

GenCore version 5.1.3  
Copyright (c) 1993 - 2002 Compugen Ltd

OM nucleic - nucleic search, using sw model

Run on: December 22, 2002, 06:13:42 ; Search time 56 Seconds  
(without alignments)  
4436.300 Million cell updates/sec

Title: US-09-691-763B-4

Sequence: 1 agcgcccttctgtagggc.....ggaaggaaggatcactt 626

Scoring table: IDENTITY\_NUC  
Gapop 10.0 , Gapext 1.0

Searched: 356696 seqs, 198428768 residues

Total number of hits satisfying chosen parameters: 713392

```
Minimum DB seq length: 0
Maximum DB seq length: 2000000000
```

Maximum DB seq length: 20000000000

post-processing:	Minimum	Match	0%
	Maximum	Match	100%

Database :

- 1: /cgn2\_6/p/odata/1/pubna/NS07\_PUBCOMB.seq.\*
- 2: /cgn2\_6/p/odata/1/pubna/PCT\_NEW\_PUB.seq.\*
- 3: /cgn2\_6/p/odata/1/pubna/NS06\_NEW\_PUB.seq.\*
- 4: /cgn2\_6/p/odata/1/pubna/NS06\_PUBCOMB.seq.\*
- 5: /cgn2\_6/p/odata/1/pubna/NS07\_NEW\_PUB.seq.\*
- 6: /cgn2\_6/p/odata/1/pubna/PCRS\_PUBCOMB.seq.\*
- 7: /cgn2\_6/p/odata/1/pubna/NS08\_NEW\_PUB.seq.\*
- 8: /cgn2\_6/p/odata/1/pubna/NS08\_PUBCOMB.seq.\*
- 9: /cgn2\_6/p/odata/1/pubna/NS09\_NEW\_PUB.seq.\*
- 10: /cgn2\_6/p/odata/1/pubna/NS09\_PUBCOMB.seq.\*
- 11: /cgn2\_6/p/odata/1/pubna/NS10\_NEW\_PUB.seq.\*
- 12: /cgn2\_6/p/odata/1/pubna/NS10\_PUBCOMB.seq.\*
- 13: /cgn2\_6/p/odata/1/pubna/NS06\_PUB.seq.\*
- 14: /cgn2\_6/p/odata/1/pubna/NS06\_PUBCOMB.seq.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Match	Query length	DB	ID	Description
1	389	62.1	811	10	US-09-925-301-278	Sequence 278, Apple
2	392	51.4	740	10	US-09-728-721-48	Sequence 48, Apple
3	392	51.4	740	10	US-09-986-617-7	Sequence 7, Apple
4	392	51.4	740	10	US-09-986-617-10	Sequence 10, Apple
5	392	51.4	740	10	US-09-841-8798-4	Sequence 4, Apple
6	392	51.4	740	10	US-09-841-8798-19	Sequence 19, Apple
7	392	51.4	740	10	US-09-925-299-648	Sequence 648, Apple
8	395	43.9	585	10	US-09-728-721-50	Sequence 50, Apple
9	275	43.9	585	10	US-09-986-617-9	Sequence 9, Apple
10	275	43.9	585	10	US-09-841-8798-6	Sequence 6, Apple
11	178.8	28.6	777	10	US-09-728-721-60	Sequence 60, Apple
12	178.8	28.6	777	10	US-09-841-8798-1	Sequence 1, Apple
13	178.8	28.6	777	10	US-09-841-8798-18	Sequence 18, Apple
14	165	26.4	579	10	US-09-728-721-62	Sequence 62, Apple
15	163.4	26.1	579	10	US-09-841-8798-3	Sequence 3, Apple
16	74	11.8	394	10	US-09-864-761-2179	Sequence 21179, Apple
17	62.8	10.0	368	10	US-09-864-761-4430	Sequence 4430, Apple
18	54.2	8.7	12425	12	US-10-023-529-50	Sequence 50, Apple
19	54.2	8.7	12425	12	US-10-023-523-50	Sequence 50, Apple

20	49	7.8	1614	12	US-10-023-529-45	Sequence 45, Appl
21	49	7.8	1614	12	US-10-023-529-45	Sequence 45, Appl
22	48	7.7	530	9	US-10-073-353-4	Sequence 4, Appl
23	48	7.7	88421	9	US-09-976-059-1	Sequence 1, Appl
24	47.6	7.6	4826	10	US-09-772-304A-1	Sequence 1, Appl
25	47	7.5	2961	10	US-09-834-975-996	Sequence 966, Appl
26	47	7.5	2961	10	US-09-834-975-997	Sequence 997, Appl
27	46.4	7.4	1506	10	US-09-963-285-9	Sequence 9, Appl
28	46.4	7.4	3289	10	US-09-963-285-8	Sequence 8, Appl
29	46.4	7.4	4156	10	US-09-963-285-3	Sequence 3, Appl
30	46.4	7.4	4366	12	US-10-004-090-668	Sequence 668, Appl
31	46.4	7.4	6456	10	US-09-963-283-1	Sequence 1, Appl
32	46	7.3	8459	10	US-09-817-913-8	Sequence 8, Appl
33	46	7.3	8459	10	US-09-817-913-8	Sequence 8, Appl
34	45.4	7.3	441	10	US-09-960-352-1021	Sequence 1021, Appl
35	44.6	7.1	1707	10	US-09-764-870-217	Sequence 217, Appl
36	44.2	7.1	1041	10	US-09-916-790-3	Sequence 3, Appl
37	44.2	7.1	1275	10	US-09-764-032-1	Sequence 1, Appl
38	44.2	7.1	1275	12	US-10-016-983-1	Sequence 1, Appl
39	44.2	7.1	1473	12	US-10-016-983-3	Sequence 3, Appl
40	44.2	7.1	2598	10	US-09-816-094-1	Sequence 1, Appl
41	44.2	7.1	2893	10	US-09-916-790-1	Sequence 1, Appl
42	44.2	7.1	7301	10	US-09-816-094-3	Sequence 3, Appl
43	44.2	7.1	7301	10	US-09-734-032-3	Sequence 3, Appl
44	44	7.0	2090	10	US-09-822-830A-105	Sequence 105, Appl
45	44	7.0	2142	10	US-09-822-687-3	Sequence 3, Appl

## ALIGNMENTS

## RESULT 1

; Sequence 278, Application US/09925301

Patent No. US20020052308A1

GENERAL INFORMATION:

APPLICANT: Rosen et al

TITLE OF INVENTION: Nucleic Acids, Proteins and Antibodies

FILE REFERENCE: PA106

CURRENT APPLICATION NUMBER: US/09/925.301

CURRENT FILING DATE: 2001-08-10

CURRENT FILING DATE: 2001-08-10  
PRIOR APPLICATION NUMBER: DGT/US00/05883

PRIOR APPLICATION NUMBER: PCF/US00/05882

PRIOR FILING DATE: 2000-03-08

;  
PRIOR APPLICATION NUMBER: 60/124,270

; PRIOR FILING DATE: 1999-03-12

; NUMBER OF SEQ ID NOS: 1694

```

; SOFTWARE: PatentIn Ver. 2.0

```

SEO ID NO 278

LENGTH: 811

TYPE: DNA

[illegible]



```
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 10
; LENGTH: 740
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-996-617-10
```

```
Query Match          51.4%; Score 322; DB 10; Length 740;
Best Local Similarity 100.0%; Pred. No. 1.2e-59;
Matches 322; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```
QY 110 CGGCTGACAGGGGTGAGCGGGGAGCGGGGATCTGAGACCATGAGGGCGCGGC 169
    |||||||
Db 734 CGGCTGACAGGGGTGAGCGGGGAGCGGGGATCTGAGACCATGAGGGCGCGGC 675
QY 170 GCGAGCCATCTGATGCGCTGAGAACCTGACCGCCGAGAGCTCAAGAGTTCAAGC 229
    |||||||
Db 674 GCGAGCCATCTGATGCGCTGAGAACCTGACCGCCGAGAGCTCAAGAGTTCAAGC 615
QY 230 TGACCTGCTGTGCGTCCGCTGCGAGGGCTAGCGGGCGATCCGCGGGGCGCGTGC 289
    |||||||
Db 614 TGACCTGCTGTGCGTCCGCTGCGAGGGCTAGCGGGCGATCCGCGGGGCGCGTGC 555
QY 290 TGTCCATGAGACGCTTGAGCTCACCAGAACCTGTGACCTTCTACCTGAGACCTAGC 349
    |||||||
Db 554 TGTCCATGAGACGCTTGAGCTCACCAGAACCTGTGACCTTCTACCTGAGACCTAGC 495
QY 350 GCGCGAGCTACCGCTGAGGTGCTGCGAGATGGGCTGACAGAGATGGCGGCGCAGC 409
    |||||||
Db 494 GCGCGAGCTACCGCTGAGGTGCTGCGAGATGGGCTGACAGAGATGGCGGCGCAGC 435
QY 410 TGCAGGCGGCACGACCAAGGG 431
    |||||||
Db 434 TGCAGGCGGCACGACCAAGGG 413
```

## RESULT 5

```
US-09-841-879B-4
; Sequence 4, Application US/09841879B
; Patent No. US20020142979A1
; GENERAL INFORMATION:
; APPLICANT: Bertin, John
; TITLE OF INVENTION: NOVEL MOLECULES OF THE CARD-RELATED PROTEIN FAMILY AND USES THERE
; FILE REFERENCE: 07334-330001
; CURRENT APPLICATION NUMBER: US/09/841,879B
; PRIORITY FILING DATE: 2001-04-24
; PRIOR FILING DATE: 2000-12-01
; PRIOR APPLICATION NUMBER: US 09/340,620
; PRIOR FILING DATE: 1999-06-28
; NUMBER OF SEQ ID NOS: 19
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 4
; LENGTH: 740
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (54)...(638)
US-09-841-879B-4
```

```
Query Match          51.4%; Score 322; DB 10; Length 740;
Best Local Similarity 100.0%; Pred. No. 1.2e-59;
Matches 322; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```
QY 110 CGGCTGACAGGGGTGAGCGGGGAGCGGGGATCTGAGACCATGAGGGCGCGGC 169
    |||||||
Db 7 CGGCTGACAGGGGTGAGCGGGGAGCGGGGATCTGAGACCATGAGGGCGCGGC 66
QY 170 GCGAGCCATCTGATGCGCTGAGAACCTGACCGCCGAGAGCTCAAGAGTTCAAGC 229
    |||||||
Db 67 GCGAGCCATCTGATGCGCTGAGAACCTGACCGCCGAGAGCTCAAGAGTTCAAGC 126
```

```
QY 230 TGAAGTGTGTGTCGTCGCGCTGACGCGAGAGGCTTACGGGGCATCCCGGGGCGCGTGC 289
    |||||||
Db 127 TGAAGTGTGTGTCGTCGCGCTGACGCGAGAGGCTTACGGGGCATCCCGGGGCGCGTGC 186
QY 290 TGTCCATGAGACGCTTGAGCTCACCAGAACCTGTGACCTTCTACTGAGAGCTTAGC 349
    |||||||
Db 187 TGTCCATGAGACGCTTGAGCTCACCAGAACCTGTGACCTTCTACTGAGAGCTTAGC 246
QY 350 GCGCGAGCTACCGCTGAGGTGCTGCGGACATGGGCGCTGACAGAGATGGCCGCGCAGC 409
    |||||||
Db 247 GCGCGAGCTACCGCTGAGGTGCTGCGGACATGGGCGCTGACAGAGATGGCCGCGCAGC 306
QY 410 TGCAGGCGGCACGACCAAGGG 431
    |||||||
Db 307 TGCAGGCGGCACGACCAAGGG 328
```

## RESULT 6

```
US-09-841-879B-19/c
; Sequence 19, Application US/09841879B
; Patent No. US20020142979A1
; GENERAL INFORMATION:
; APPLICANT: Bertin, John
; TITLE OF INVENTION: NOVEL MOLECULES OF THE CARD-RELATED PROTEIN FAMILY AND USES TH
; FILE REFERENCE: 07334-330001
; CURRENT APPLICATION NUMBER: US/09/841,879B
; PRIORITY FILING DATE: 2001-04-24
; PRIOR APPLICATION NUMBER: US 09/728,721
; PRIOR FILING DATE: 2000-12-01
; PRIOR APPLICATION NUMBER: US 09/340,620
; PRIOR FILING DATE: 1999-06-28
; NUMBER OF SEQ ID NOS: 19
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 19
; LENGTH: 740
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-841-879B-19
```

```
Query Match          51.4%; Score 322; DB 10; Length 740;
Best Local Similarity 100.0%; Pred. No. 1.2e-59;
Matches 322; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```
QY 110 CGGCTGACAGGGGTGAGCGGGGAGCGGGGATCTGAGACCATGAGGGCGCGGC 169
    |||||||
Db 734 CGGCTGACAGGGGTGAGCGGGGAGCGGGGATCTGAGACCATGAGGGCGCGGC 675
QY 170 GCGAGCCATCTGATGCGCTGAGAACCTGACCGCCGAGAGCTCAAGAGTTCAAGC 229
    |||||||
Db 674 GCGAGCCATCTGATGCGCTGAGAACCTGACCGCCGAGAGCTCAAGAGTTCAAGC 615
QY 230 TGAAGTGTGTGTCGTCGCGCTGACGCGAGAGGCTTACGGGGCATCCCGGGGCGCGTGC 289
    |||||||
Db 614 TGAAGTGTGTGTCGTCGCGCTGACGCGAGAGGCTTACGGGGCATCCCGGGGCGCGTGC 555
QY 290 TGTCCATGAGACGCTTGAGCTCACCAGAACCTGTGACCTTCTACTGAGAGCTTAGC 349
    |||||||
Db 554 TGTCCATGAGACGCTTGAGCTCACCAGAACCTGTGACCTTCTACTGAGAGCTTAGC 495
QY 350 GCGCGAGCTACCGCTGAGGTGCTGCGGACATGGGCGCTGACAGAGATGGCCGCGCAGC 409
    |||||||
Db 494 GCGCGAGCTACCGCTGAGGTGCTGCGGACATGGGCGCTGACAGAGATGGCCGCGCAGC 435
QY 410 TGCAGGCGGCACGACCAAGGG 431
    |||||||
Db 434 TGCAGGCGGCACGACCAAGGG 413
```

## RESULT 7

```
US-09-925-299-648
; Sequence 648, Application US/09925299
; Patent No. US20020055627A1
; GENERAL INFORMATION:
```

APPLICANT: Rosen et al.  
TITLE OF INVENTION: Nucleic Acids, Proteins and Antibodies  
FILE REFERENCE: PA102  
CURRENT APPLICATION NUMBER: US/09/925,299  
CURRENT FILING DATE: 2001-08-10  
PRIOR APPLICATION NUMBER: PCT/US00/05883  
PRIOR FILING DATE: 2000-03-08  
PRIOR APPLICATION NUMBER: 60/124,270  
PRIOR FILING DATE: 1999-03-12  
NUMBER OF SEQ ID NOS: 1556  
SOFTWARE: PatentIn Ver. 2.0  
SEQ ID NO: 648  
LENGTH: 432  
TYPE: DNA  
ORGANISM: Homo sapiens  
US-09-925-299-648

Query Match 51.1%; Score 320; DB 10; Length 432;  
Best Local Similarity 100.0%; Pred. No. 3,1e-59;  
Matches 320; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 112 GCTCAGCGGGGTAGCGGGGACGCGCGGGAGCTTGAGACCATGGGGCGCGCGC 171  
DB 8 GCTCAGCGGGGTAGCGGGGACGCGCGGGAGCTTGAGACCATGGGGCGCGCGC 67  
QY 172 GACGCCATCTCTGATGCTGAGAACTGACCGCGAGAGAGCTCAAGAATTCAAGCTG 231  
DB 68 GACGCCATCTCTGATGCTGAGAACTGACCGCGAGAGAGCTCAAGAATTCAAGCTG 127  
QY 232 AAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 291  
DB 128 AAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 187  
QY 292 TCCATGAGACCCCTTGAGACCTGACGAAAGTGTGCTTACTGAGAACTACAGGC 351  
DB 188 TCCATGAGACCCCTTGAGACCTGACGAAAGTGTGCTTACTGAGAACTACAGGC 247  
QY 352 GCCGAGCTCAACGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 411  
DB 248 GCCGAGCTCAACGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 307  
QY 412 CAGCGCGCGCACGACCAAGG 431  
DB 308 CAGCGCGCGCACGACCAAGG 327

RESULT 8  
US-09-728-721-50  
Sequence 50, Application US/09728721  
Patent No. US20020061845A1  
GENERAL INFORMATION:

APPLICANT: Bertin, John  
TITLE OF INVENTION: NOVEL MOLECULES OF THE CARD-RELATED PROTEIN FAMILY AND USES THERE  
FILE REFERENCE: 07334-124001  
CURRENT APPLICATION NUMBER: US/09/728,721  
CURRENT FILING DATE: 2000-12-01  
PRIOR APPLICATION NUMBER: 09/340,620  
PRIOR FILING DATE: 1999-06-28  
PRIOR APPLICATION NUMBER: US 09/207,359  
PRIOR FILING DATE: 1998-12-08  
PRIOR APPLICATION NUMBER: US 09/099,041  
PRIOR FILING DATE: 1998-06-17  
PRIOR APPLICATION NUMBER: 09/019,942  
PRIOR FILING DATE: 1998-02-06  
NUMBER OF SEQ ID NOS: 71  
SOFTWARE: FastSeq for Windows Version 4.0  
SEQ ID NO: 50  
LENGTH: 585  
TYPE: DNA  
ORGANISM: Homo sapiens  
US-09-728-721-50

Query Match 43.9%; Score 275; DB 10; Length 585;

Best Local Similarity 100.0%; Pred. No. 9.4e-50;  
Matches 275; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 157 ATGGGCGCGCGCGGACCCATCTGATGCGCTGAGAACTGACCGCGGAGAGCTC 216  
DB 1 ATGGGCGCGCGCGGACCCATCTGATGCGCTGAGAACTGACCGCGGAGAGCTC 60  
QY 217 AAGAATTCAAGCTGAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 276  
DB 61 AAGAATTCAAGCTGAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 120  
QY 277 CGGGGCGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 336  
DB 121 CGGGGCGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 180  
QY 337 CTGAGAGCTTACGCGCGGAGCTTACCGCTTACCGCTTACCGCTTACCGCTTAC 396  
DB 181 CTGAGAGCTTACGCGCGGAGCTTACCGCTTACCGCTTACCGCTTACCGCTTAC 240  
QY 397 ATGGCGGCGAGCTGACGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 431  
DB 241 ATGGCGGCGAGCTGACGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 275

RESULT 9  
US-09-996-617-9  
Sequence 9, Application US/09996617  
Patent No. US20020128198A1  
GENERAL INFORMATION:  
APPLICANT: Bertin, John  
TITLE OF INVENTION: NOVEL MOLECULES OF THE CARD-RELATED  
FILE REFERENCE: 07334-340001  
CURRENT APPLICATION NUMBER: US/09/996,617  
CURRENT FILING DATE: 2001-11-27  
PRIOR APPLICATION NUMBER: 09/931,071  
PRIOR FILING DATE: 2001-08-15  
PRIOR APPLICATION NUMBER: 09/428,252  
PRIOR FILING DATE: 1999-10-27  
PRIOR APPLICATION NUMBER: 09/340,620  
PRIOR FILING DATE: 1999-06-28  
NUMBER OF SEQ ID NOS: 10  
SOFTWARE: FastSeq for Windows Version 4.0  
SEQ ID NO: 9  
LENGTH: 585  
TYPE: DNA  
ORGANISM: Homo sapiens  
US-09-996-617-9

Query Match 43.9%; Score 275; DB 10; Length 585;  
Best Local Similarity 100.0%; Pred. No. 9.4e-50;  
Matches 275; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 157 ATGGGCGCGCGCGGACCCATCTGATGCGCTGAGAACTGACCGCGGAGAGCTC 216  
DB 1 ATGGGCGCGCGCGGACCCATCTGATGCGCTGAGAACTGACCGCGGAGAGCTC 60  
QY 217 AAGAATTCAAGCTGAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 276  
DB 61 AAGAATTCAAGCTGAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 120  
QY 277 CGGGGCGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 336  
DB 121 CGGGGCGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 180  
QY 337 CTGAGAGCTTACGCGCGGAGCTTACCGCTTACCGCTTACCGCTTACCGCTTAC 396  
DB 181 CTGAGAGCTTACGCGCGGAGCTTACCGCTTACCGCTTACCGCTTACCGCTTAC 240  
QY 397 ATGGCGGCGAGCTGACGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 431  
DB 241 ATGGCGGCGAGCTGACGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 275

```
RESULT 10
US-09-841-879b-6
; Sequence 6, Application US/09841879B
; Patent No. US20020142979A1
; GENERAL INFORMATION:
; APPLICANT: Bertlin, John
; TITLE OF INVENTION: NOVEL MOLECULES OF THE CARD-RELATED PROTEIN FAMILY AND USES THERE
; FILE REFERENCE: 07334-330001
; CURRENT APPLICATION NUMBER: US/09/841,879B
; CURRENT FILING DATE: 2001-04-24
; PRIOR APPLICATION NUMBER: US 09/728,721
; PRIOR FILING DATE: 2000-12-01
; PRIOR APPLICATION NUMBER: US 09/340,620
; PRIOR FILING DATE: 1999-06-28
; NUMBER OF SEQ ID NOS: 19
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 6
; LENGTH: 585
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-841-879b-6

Query Match          43.9%; Score 275; DB 10; Length 585;
Best Local Similarity 100.0%; Pred. No. 9,4e-50;
Matches 275; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 157 ATGGGGGGGGGGGGGAGCGCATCTGGATGGGCTGGAGAACTGACCGCGGAGAGCTC 216
DB 1 ATGGGGGGGGGGGGGAGCGCATCTGGATGGGCTGGAGAACTGACCGCGGAGAGCTC 60
QY 217 AAGAAGTTCAGACTGAAGCTGCTGTGCGTCCGCTGGCGGAGGGCTACAGGGCCATCCCG 276
DB 61 AAGAAGTTCAGACTGAAGCTGCTGTGCGTCCGCTGGCGGAGGGCTACAGGGCCATCCCG 120
QY 277 CGGGGGGGGGGGCTGTCTCATGAGGAGCGCTTGGAGCTCAGCGAAGCTGGTCACTTCTAC 336
DB 121 CGGGGGGGGGGGCTGTCTCATGAGGAGCGCTTGGAGCTCAGCGAAGCTGGTCACTTCTAC 180
QY 337 CTGGAGACCTAGCGGGCGGAGCTCACCCTTAACGTGTGCGCGAGCATGGGCTTGCAGAG 396
DB 181 CTGGAGACCTAGCGGGCGGAGCTCACCCTTAACGTGTGCGCGAGCATGGGCTTGCAGAG 240
QY 397 ATGGCGGGGAGCTGCGAGGGGGCGGACGACGACGAGG 431
DB 241 ATGGCGGGGAGCTGCGAGGGGGCGGACGACGACGAGG 275

RESULT 11
US-09-728-721-60
; Sequence 60, Application US/09728721
; Patent No. US20020061845A1
; GENERAL INFORMATION:
; APPLICANT: Bertlin, John
; TITLE OF INVENTION: NOVEL MOLECULES OF THE CARD-RELATED PROTEIN FAMILY AND USES THERE
; FILE REFERENCE: 07334-124001
; CURRENT APPLICATION NUMBER: US/09/728,721
; CURRENT FILING DATE: 2000-12-01
; PRIOR APPLICATION NUMBER: 09/340,620
; PRIOR FILING DATE: 1999-06-28
; PRIOR APPLICATION NUMBER: US 09/207,359
; PRIOR FILING DATE: 1998-12-08
; PRIOR APPLICATION NUMBER: US 09/099,041
; PRIOR FILING DATE: 1998-06-17
; PRIOR APPLICATION NUMBER: US 09/019,942
; PRIOR FILING DATE: 1998-02-06
; NUMBER OF SEQ ID NOS: 71
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 60
; LENGTH: 777
; TYPE: DNA
; ORGANISM: Mus musculus
; FEATURE:
```

```
; NAME/KEY: CDS
; LOCATION: (89)...(667)
US-09-728-721-60

Query Match          28.6%; Score 178.8; DB 10; Length 777;
Best Local Similarity 70.9%; Pred. No. 1.7e-29;
Matches 251; Conservative 0; Mismatches 102; Indels 1; Gaps 1;

QY 81 CGCCCGGAGCCGACTCTCTCTGCTGGCGGCTGAGC-GGGGTGAGCGCGGACCGG 139
DB 12 GCTCCGGAGCAGCAGCTGCTGAGCGAGCAGCAGCAGCAAGATGAAGTGAACCGCGG 71
QY 140 CCGGGGATCCTGGAGCAGTGGGGGCGGCGGCGAGCCATCTGAGTGGCTGGGAACC 199
DB 72 CTGCCACCCCGAGAGCAGTGGGGCGGCGGCGAGAGTGCATCTGAGACCTCTTAAACT 131
QY 200 TGACCGCCGAGGAGCTCAAGAAAGTTCAAGCTGAAGCTGTGCGCTGGCGGAGG 259
DB 132 TGTGAGGGGATGAACCTCAAAAGTTCAAGATGAAGCTGTGACAGTGCACCTGGAGAG 191
QY 260 GCTACGGGGGATCCCGGGGGGCGGCTGTCTCATGAGAGCGCTTGGACCTCAGCAGCA 319
DB 192 GCTATGGGCGCATCCAGCCGCGGCGCTGCTGAGATGAGACGCTATGATCTACTGAC 251
QY 320 AGCTGTGACTTCTACTGTGAGACTTACGCGCGGAGCTCACCCTTAACCTGCTGCGG 379
DB 252 AACTTGTACACTACTTCTGTGAGTGTGTATGCTTGGAGCTCAACATGACTGTGCTTAG 311
QY 380 ACATGGGCTGTGAGAGATGGCGGGGAGCTGCAGCGGCGCAGCCAGGAGTG 433
DB 312 ACATGGCTTACAGGAGCTGTGCTGAGAGCTGCACAAAGCTAAAGAGAGTCTG 365

RESULT 12
US-09-841-879b-1
; Sequence 1, Application US/09841879B
; Patent No. US20020142979A1
; GENERAL INFORMATION:
; APPLICANT: Bertlin, John
; TITLE OF INVENTION: NOVEL MOLECULES OF THE CARD-RELATED PROTEIN FAMILY AND USES TH
; FILE REFERENCE: 07334-330001
; CURRENT APPLICATION NUMBER: US/09/841,879B
; CURRENT FILING DATE: 2001-04-24
; PRIOR APPLICATION NUMBER: US 09/728,721
; PRIOR FILING DATE: 2000-12-01
; PRIOR APPLICATION NUMBER: US 09/340,620
; PRIOR FILING DATE: 1999-06-28
; NUMBER OF SEQ ID NOS: 19
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 1
; LENGTH: 777
; TYPE: DNA
; ORGANISM: Mus musculus
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (89)...(667)
US-09-841-879b-1

Query Match          28.6%; Score 178.8; DB 10; Length 777;
Best Local Similarity 70.9%; Pred. No. 1.7e-29;
Matches 251; Conservative 0; Mismatches 102; Indels 1; Gaps 1;

QY 81 CGCCCGGAGCCGACTCTCTCTGCTGGCGGCTGAGC-GGGGTGAGCGCGGACCGG 139
DB 12 GCTCCGGAGCAGCAGCTGCTGAGCGAGCAGCAGCAGCAAGATGAAGTGAACCGCGG 71
QY 140 CCGGGGATCCTGGAGCAGTGGGGGCGGCGGCGAGCCATCTGAGTGGCTGGGAACC 199
DB 72 CTGCCACCCCGAGAGCAGTGGGGCGGCGGCGAGAGTGCATCTGAGACCTCTTAAACT 131
QY 200 TGACCGCCGAGGAGCTCAAGAAAGTTCAAGCTGAAGCTGTGCGCTGGCGGAGG 259
DB 132 TGTGAGGGGATGAACCTCAAAAGTTCAAGATGAAGCTGTGACAGTGCACCTGGAGAG 191
```









GenCore version 5.1.3  
Copyright (c) 1993 - 2002 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: December 22, 2002, 05:21:46 ; Search time 2672 Seconds  
(without alignments)

6818.247 Million cell updates/sec

Title: US-09-691-763b-4

Sequence: 1 aggccttctgctggaagc.....ggaaggaaggaagcactt 626

Scoring table: IDENTITY\_NUC  
Gapop 10.0 , Gapext 1.0

Searched: 2054640 seqs, 14551402878 residues

Total number of hits satisfying chosen parameters: 4109280

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

GenBank: 1: gb\_da: 2: gb\_hg: 3: gb\_in: 4: gb\_om: 5: gb\_ov: 6: gb\_pat: 7: gb\_pi: 8: gb\_pi: 9: gb\_pi: 10: gb\_pi: 11: gb\_pi: 12: gb\_pi: 13: gb\_pi: 14: gb\_pi: 15: gb\_pi: 16: gb\_pi: 17: gb\_pi: 18: gb\_pi: 19: gb\_pi: 20: gb\_pi: 21: gb\_pi: 22: gb\_pi: 23: gb\_pi: 24: gb\_pi: 25: gb\_pi: 26: gb\_pi: 27: gb\_pi: 28: gb\_pi: 29: gb\_pi: 30: gb\_pi: 31: gb\_pi: 32: gb\_pi: 33: gb\_pi: 34: gb\_pi: 35: gb\_pi: 36: gb\_pi: 37: gb\_pi: 38: gb\_pi: 39: gb\_pi: 40: gb\_pi: 41: gb\_pi:

Pred. No. is the number of results predicted by chance to have a

score greater than or equal to the score of the result being printed,  
and is derived by analysis of the total score distribution.

# SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	626	100.0	626	AX118621	AX118621 Sequence
2	626	100.0	2821	AX118618	AX118618 Sequence
3	626	100.0	2821	AF184072	AF184072 Homo sapi
4	626	100.0	23305	AC009088	AC009088 Homo sapi
5	368	58.8	779	AX012720	AX012720 Sequence
6	361	57.7	782	AX459863	AX459863 Sequence
7	361	57.7	782	AB023416	AB023416 Homo sapi
8	342.8	54.8	713	AF186416	AF186416 Sequence
9	342.8	54.8	713	AF255794	AF255794 Homo sapi
10	342.6	54.7	405	AX118643	AX118643 Sequence
11	342.6	54.7	770	AX118619	AX118619 Sequence
12	342.6	54.7	770	AF184073	AF184073 Homo sapi
13	340	54.3	340	AX118622	AX118622 Sequence
14	322	51.4	740	AX082246	AX082246 Sequence
15	322	51.4	740	AF384665	AF384665 Homo sapi
16	301	48.1	740	AK000211	AK000211 Homo sapi
17	275	43.9	585	AX082248	AX082248 Sequence
18	275	43.9	712	AF310103	AF310103 Homo sapi
19	268	42.8	772	BC013569	BC013569 Homo sapi
20	232.4	37.1	768	BC004470	BC004470 Homo sapi
21	179	28.6	1115	AB059327	AB059327 Mus muscu
22	179	28.6	7117	AC101465	AC101465 Mus muscu
23	179	28.6	19128	AC124461	AC124461 Mus muscu
24	179	28.6	225984	AC093175	AC093175 Mus muscu
25	178.8	28.6	777	AX082258	AX082258 Sequence
26	178.6	28.5	732	BC008252	BC008252 Mus muscu
27	178.6	28.5	770	AB032249	AB032249 Mus muscu
28	178.6	28.5	803	AX118637	AX118637 Sequence
29	177.6	28.4	551	AX459859	AX459859 Sequence
30	175.6	28.1	701	AF310104	AF310104 Mus muscu
31	174.4	27.9	556	AF086332	AF086332 Homo sapi
32	174.4	27.9	587	AF454669	AF454669 Homo sapi
33	174.4	27.9	807	AX048091	AX048091 Sequence
34	169.4	27.1	270	AX417218	AX417218 Sequence
35	166	26.5	168975	AC117170	AC117170 Rattus no
36	165	26.4	579	AX082260	AX082260 Sequence
37	129	20.6	52521	AC106629	AC106629 Rattus no
38	117.8	18.8	605	AX118639	AX118639 Sequence
39	86.2	13.8	16891	AF111163	AF111163 Homo sapi
40	86.2	13.8	168700	AC009115	AC009115 Homo sapi
41	86.2	13.8	239566	HSARJ03147	HSARJ03147 Homo sapi
42	74.8	11.9	2825	HSARENO	Y14441 Homo sapien
43	74.8	11.9	3500	AX459861	AX459861 Sequence
44	74.8	11.9	3511	AF018080	AF018080 Homo sapi
45	73	11.7	95038	AF015416	AF015416 Homo sapi

# ALIGNMENTS

RESULT 1	AX118621	626 bp	DNA	linear	PAT 11-MAY-2001
LOCUS	AX118621				
DEFINITION	Sequence 4 from Patent WO0129235.				
ACCESSION	AX118621				
VERSION	AX118621.1	GI:14035572			
KEYWORDS					
SOURCE	human.				
ORGANISM	Homo sapiens				
REFERENCE	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;				
AUTHORS	Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.				
TITLE	Vertino, P.M.				
JOURNAL	1 (bases 1 to 626)				
	Patent: WO 0129235-A 4 26-APR-2001;				

FEATURES Emory University (US)  
Location/Qualifiers  
SOURCE 1. 626  
/Organism="Homo sapiens"  
/db\_xref="taxon:9606"  
BASE COUNT 103 a 216 c 216 g 91 t  
ORIGIN

Query Match 100.0%; Score 626; DB 6; Length 626;  
Best Local Similarity 100.0%; Pred. No. 1.8e-75;  
Matches 626; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 AGCGCTTTTCTGAGAGGCAACGACCGGCGGAGCTCGAGACCAAGTGGAGG 60  
DB 1 AGCGCTTTTCTGAGAGGCAACGACCGGCGGAGCTCGAGACCAAGTGGAGG 60  
QY 61 AAGCGGGAGATCCAGGTTCCGCCCGGAGCGACTTCTCTGTGGGCTGACAG 120  
DB 61 AAGCGGGAGATCCAGGTTCCGCCCGGAGCGACTTCTCTGTGGGCTGACAG 120  
QY 121 GGGTGAAGCGGCGGAGCGGCGGAGCTCTGAGCCATGGGCGCGCGAGCCATC 180  
DB 121 GGGTGAAGCGGCGGAGCGGCGGAGCTCTGAGCCATGGGCGCGCGAGCCATC 180  
QY 181 CTGATGCGCTGGAGAACTGACCGCCGAGAGGCTCAAGAAATTCAAGCTGAAGCTGCTG 240  
DB 181 CTGATGCGCTGGAGAACTGACCGCCGAGAGGCTCAAGAAATTCAAGCTGAAGCTGCTG 240  
QY 241 TCGGTGCGCGCGGCGGAGGCTACGGGCGCATCCCGGGGGCGGCTGCTGCTGATGAGC 300  
DB 241 TCGGTGCGCGCGGCGGAGGCTACGGGCGCATCCCGGGGGCGGCTGCTGCTGATGAGC 300  
QY 301 GCCTTGAGCTTCAACGACAAAGCTGTGCTTCACTTGAAGAGCTACGGGCGCGAGCTC 360  
DB 301 GCCTTGAGCTTCAACGACAAAGCTGTGCTTCACTTGAAGAGCTACGGGCGCGAGCTC 360  
QY 361 ACCGCTTAAGCTGCTGGCGGACATGGGCGCTGAGAGATGGCCGGGAGCTGAGAGCGGCC 420  
DB 361 ACCGCTTAAGCTGCTGGCGGACATGGGCGCTGAGAGATGGCCGGGAGCTGAGAGCGGCC 420  
QY 421 ACAGCAGAGGATGAGCGCGCCCGCTTCCCTCCACCGCGTCTTCCCTCCACCGACAC 480  
DB 421 ACAGCAGAGGATGAGCGCGCCCGCTTCCCTCCACCGCGTCTTCCCTCCACCGACAC 480  
QY 481 AGCGCTTACCGCGCGGCTTCCCTTCTGTTCTCTACCCCTTAACAAGTGTCTC 540  
DB 481 AGCGCTTACCGCGCGGCTTCCCTTCTGTTCTCTACCCCTTAACAAGTGTCTC 540  
QY 541 TACCGGAAGAGAGGCTCCCGACGCTTGGCTACCGACCAAGGAGCCCGCGCCACAGGC 600  
DB 541 TACCGGAAGAGAGGCTCCCGACGCTTGGCTACCGACCAAGGAGCCCGCGCCACAGGC 600  
QY 601 GGGAGGGAAGGGAAGGGATCACTT 626  
DB 601 GGGAGGGAAGGGAAGGGATCACTT 626

RESULT 2  
AX118618 2821 bp DNA 1linear PAT 11-MAY-2001  
LOCUS AX118618  
DEFINITION Sequence 1 from Patent WO0129235.  
ACCESSION AX118618  
VERSION AX118618.1 GI:14035569  
KEYWORDS  
SOURCE human.  
ORGANISM Homo sapiens.  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.  
REFERENCE 1 (bases 1 to 2821)  
AUTHORS Vertino,P.M.  
Tms1 compositions and methods of use  
JOURNAL Patent: WO 0129235-A 1 26-APR-2001;  
Emory University (US)

FEATURES Emory University (US)  
Location/Qualifiers  
SOURCE 1. 2821  
/Organism="Homo sapiens"  
/db\_xref="taxon:9606"  
BASE COUNT 614 a 864 c 757 g 586 t  
ORIGIN

Query Match 100.0%; Score 626; DB 6; Length 2821;  
Best Local Similarity 100.0%; Pred. No. 1.2e-75;  
Matches 626; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 AGCGCTTTTCTGAGAGGCAACGACCGGCGGAGCTCGAGACCAAGTGGAGG 60  
DB 1100 AGCGCTTTTCTGAGAGGCAACGACCGGCGGAGCTCGAGACCAAGTGGAGG 1159  
QY 61 AAGCGGGAGATCCAGGTTCCGCCCGGAGCGACTTCTCTGTGGGCTGACAG 120  
DB 1160 AAGCGGGAGATCCAGGTTCCGCCCGGAGCGACTTCTCTGTGGGCTGACAG 1219  
QY 121 GGGTGAAGCGGCGGAGCGGCGGAGCTCTGAGCCATGGGCGCGCGAGCCATC 180  
DB 1220 GGGTGAAGCGGCGGAGCGGCGGAGCTCTGAGCCATGGGCGCGCGAGCCATC 1279  
QY 181 CTGATGCGCTGGAGAACTGACCGCCGAGAGGCTCAAGAAATTCAAGCTGAAGCTGCTG 240  
DB 1280 CTGATGCGCTGGAGAACTGACCGCCGAGAGGCTCAAGAAATTCAAGCTGAAGCTGCTG 1339  
QY 241 TCGGTGCGCGCGGCGGAGGCTACGGGCGCATCCCGGGGGCGGCTGCTGCTGATGAGC 300  
DB 1340 TCGGTGCGCGCGGCGGAGGCTACGGGCGCATCCCGGGGGCGGCTGCTGCTGATGAGC 1399  
QY 301 GCCTTGAGCTTCAACGACAAAGCTGTGCTTCACTTGAAGAGCTACGGGCGCGAGCTC 360  
DB 1400 GCCTTGAGCTTCAACGACAAAGCTGTGCTTCACTTGAAGAGCTACGGGCGCGAGCTC 1459  
QY 361 ACCGCTTAAGCTGCTGGCGGACATGGGCGCTGAGAGATGGCCGGGAGCTGAGAGCGGCC 420  
DB 1460 ACCGCTTAAGCTGCTGGCGGACATGGGCGCTGAGAGATGGCCGGGAGCTGAGAGCGGCC 1519  
QY 421 ACAGCAGAGGATGAGCGCGCCCGCTTCCCTCCACCGCGTCTTCCCTCCACCGACAC 480  
DB 1520 ACAGCAGAGGATGAGCGCGCCCGCTTCCCTCCACCGCGTCTTCCCTCCACCGACAC 1579  
QY 481 AGCGCTTACCGCGCGGCTTCCCTTCTGTTCTCTACCCCTTAACAAGTGTCTC 540  
DB 1580 AGCGCTTACCGCGCGGCTTCCCTTCTGTTCTCTACCCCTTAACAAGTGTCTC 1639  
QY 541 TACCGGAAGAGAGGCTCCCGACGCTTGGCTACCGACCAAGGAGCCCGCGCCACAGGC 600  
DB 1640 TACCGGAAGAGAGGCTCCCGACGCTTGGCTACCGACCAAGGAGCCCGCGCCACAGGC 1699  
QY 601 GGGAGGGAAGGGAAGGGATCACTT 626  
DB 1700 GGGAGGGAAGGGAAGGGATCACTT 1725

RESULT 3  
AF184072 2821 bp DNA 1linear PRI 02-FEB-2001  
LOCUS AF184072  
DEFINITION Homo sapiens target of methylation-induced silencing 1 (TMS1) gene,  
complete cds.  
ACCESSION AF184072  
VERSION AF184072.1 GI:9863861  
KEYWORDS  
SOURCE human.  
ORGANISM Homo sapiens.  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.  
REFERENCE 1 (bases 1 to 2821)  
AUTHORS Conway,K.E., McConnell,B.B., Bowring,C.E., Donald,C.D., Warren,S.T.  
and Vertino,P.M.  
TMS1, a novel proapoptotic caspase recruitment domain protein, is a  
target of methylation-induced gene silencing in human breast

JOURNAL Cancer Res. 60 (22), 6236-6242 (2000)  
MEDLINE 20552139  
PUBMED 11107776  
REFERENCE 2 (bases 1 to 2821)  
AUTHORS McConnell,B.B. and Vertino,P.M.  
TITLE Activation of a caspase-9-mediated apoptotic pathway by subcellular redistribution of the novel caspase recruitment domain protein TMS1  
JOURNAL Cancer Res. 60 (22), 6243-6247 (2000)  
MEDLINE 20552140  
PUBMED 11103777  
REFERENCE 3 (bases 1 to 2821)  
AUTHORS Vertino,P.M.  
TITLE Direct Submission  
JOURNAL Submitted (09-SEP-1999) Radiation Oncology, Emory University School of Medicine, 145 Edgewood Avenue, SE, Atlanta, GA 30335, USA  
FEATURES  
source 1..2821  
/organism="Homo sapiens"  
/db\_xref="taxon:9606"  
/cell\_type="sperm"  
/clone\_lib="Research Genetics CTRB-HSP-C BAC library"  
repeat\_region complement(44..93)  
/note="LINE"  
/rpl\_family="L2"  
/rpl\_type-dispersed complement(95..256)  
/note="SINE"  
repeat\_region /rpl\_family="Alu"  
/rpl\_type-dispersed complement(257..408)  
/note="LINE"  
/rpl\_family="L2"  
/rpl\_type-dispersed 1111..1722  
/note="CpG island"  
misc\_feature 1199..2460  
/gene="TMS1"  
join(1190..1529,1743..1799,2105..2460)  
/gene="TMS1"  
/product="target of methylation-induced silencing 1"  
exon 1190..1529  
/gene="TMS1"  
/number=1  
join(1256..1529,1743..1799,2105..2361)  
/gene="TMS1"  
/function="Involved in apoptosis"  
/note="contains caspase-recruitment domain; CARD protein; target of methylation-mediated gene silencing in Homo sapien tumors"  
codon\_start=1  
/product="target of methylation-induced silencing 1"  
protein\_id="AAG01187.1"  
/db\_xref="GI:9863862"  
/translation="MGRARDATLALLENLTAEIKKFKKLLSVLRSGYGRIRGAL  
LSMDALDTKIVSFLEYGAEITANVLRMGLOEAGOLQATHQSGAAPAGIA  
PPOSAAPKGLHFIIDHRAALILRVINEMLDALYGVITDEQYQAVRAEPTNSKM  
KLSEFTPAWNTCKDILLQALRESQSYVEDLERS"  
exon 1743..1799  
/gene="TMS1"  
/number=2  
exon 2105..2460  
/gene="TMS1"  
/number=3  
polya\_signal 2438..2443  
/gene="TMS1"  
BASE COUNT 614 a 864 c 757 g 586 t  
ORIGIN  
Query Match 100.0%; Score 626; DB 9; Length 2821;  
Best Local Similarity 100.0%; Pred. No. 1.2e-75;  
Matches 626; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 AGCGCTTTTGTGAGAGGCAACGACCGGGGAGTTCGGAGACAGAGTGGAGG 60  
|||||  
Db 1100 AGCGCTTTTGTGAGAGGCAACGACCGGGGAGTTCGGAGACAGAGTGGAGG 1159  
QY 61 AAGCGGGGAGTCCAGTTCGCCGCCGAGCCGACCTTCTCTGTCGGGGGTCGACGC 120  
|||||  
Db 1160 AAGCGGGGAGTCCAGTTCGCCGCCGAGCCGACCTTCTCTGTCGGGGGTCGACGC 1219  
QY 121 GGGTACCGGGGAGTCCAGTTCGCCGCCGAGCCGACCTTCTCTGTCGGGGGTCGACGC 180  
|||||  
Db 1220 GGGTACCGGGGAGTCCAGTTCGCCGCCGAGCCGACCTTCTCTGTCGGGGGTCGACGC 1279  
QY 181 CTGATGCGCTGAGAGACTGACCGCCGAGAGAGCTCAAGAGTTCAAGTCAAGTCTGTG 240  
|||||  
Db 1280 CTGATGCGCTGAGAGACTGACCGCCGAGAGAGCTCAAGAGTTCAAGTCAAGTCTGTG 1339  
QY 241 TCGTTCGCGTGGCGAGAGGCTACCGGGGCGCATCCCGGGGCGCGCTGCTGCATGAGAC 300  
|||||  
Db 1340 TCGTTCGCGTGGCGAGAGGCTACCGGGGCGCATCCCGGGGCGCGCTGCTGCATGAGAC 1399  
QY 301 GCCTTGGACCTCACCCAGCAAGCTGTGCTTCTACTGAGAGCTACGCGCGGAGAGCTC 360  
|||||  
Db 1400 GCCTTGGACCTCACCCAGCAAGCTGTGCTTCTACTGAGAGCTACGCGCGGAGAGCTC 1459  
QY 361 ACCGCTAACGTCGCGGACATGGGCTTCGACGAGATGGCCGCGGACGTGACGCGGC 420  
|||||  
Db 1460 ACCGCTAACGTCGCGGACATGGGCTTCGACGAGATGGCCGCGGACGTGACGCGGC 1519  
QY 421 ACGACCAAGGATGAGCGCGCGCGCTTCCCTCCACCGCTTCCCTCCACCGACAC 480  
|||||  
Db 1520 ACGACCAAGGATGAGCGCGCGCGCTTCCCTCCACCGCTTCCCTCCACCGACAC 1579  
QY 481 AGCGTTACCCCGCGGCGCTTCCCTCCCTTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 540  
|||||  
Db 1580 AGCGTTACCCCGCGGCGCTTCCCTCCCTTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 1639  
QY 541 TACCGGAAGGAGAGTCCCGACGCTTGGCTTGGCTTGGCTTGGCTTGGCTTGGCTTGG 600  
|||||  
Db 1640 TACCGGAAGGAGAGTCCCGACGCTTGGCTTGGCTTGGCTTGGCTTGGCTTGGCTTGG 1699  
QY 601 GCGAAGGAGGAGGAGGAGGATCATCTT 626  
|||||  
Db 1700 GCGAAGGAGGAGGAGGAGGATCATCTT 1725  
RESULT 4  
AC009088/c 233305 bp DNA linear PRI 03-APR-2002  
LOCUS AC009088  
DEFINITION Homo sapiens chromosome 16 clone RP11-388M20, complete sequence.  
ACCESSION AC009088  
VERSION AC009088.7 GI:19909387  
KEYWORDS HNG.  
SOURCE Homo sapiens.  
ORGANISM Homo sapiens.  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.  
REFERENCE 1 (bases 1 to 233305)  
AUTHORS DOE Joint Genome Institute.  
TITLE Sequencing of Human Chromosome 16  
JOURNAL Unpublished  
REFERENCE 2 (bases 1 to 233305)  
AUTHORS DOE Joint Genome Institute.  
TITLE Direct Submission  
JOURNAL Submitted (03-AUG-1999) Production Sequencing Facility, DOE Joint  
Genome Institute, 2800 Mitchell Drive, Walnut Creek, CA 94598, USA  
REFERENCE 3 (bases 1 to 233305)  
AUTHORS DOE Joint Genome Institute.  
TITLE Direct Submission  
JOURNAL Submitted (03-APR-2002) Production Sequencing Facility, DOE Joint  
Genome Institute, 2800 Mitchell Drive, Walnut Creek, CA 94598, USA  
COMMENT On Apr 3, 2002 this sequence version replaced gi:16924080.  
Sequence Quality Assessment:  
This entry has been annotated with sequence quality

estimates computed by the Phrap assembly program.  
All manually edited bases have been reduced to quality zero.  
Quality levels above 40 are expected to have less than  
1 error in 10,000 bp.  
Base-by-base quality values are not generally visible from the  
Genbank flat file format but are available as part  
of this entry's ASN.1 file.

#### FEATURES

source 1. 233305  
/organism="Homo sapiens"  
/db\_xref="taxon:9606"  
/chromosome="16"  
/clone="RP11-388M20"  
misc\_feature 1. 233305  
/note="clone overlap not submitted"

BASE COUNT 54492 a 59306 c 61380 g 58127 t  
ORIGIN

Query Match 100.0%; Score 626; DB 9; Length 233305;  
Best Local Similarity 100.0%; Pred. No. 4e-76;  
Matches 626; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 AGGCGCTTTGCTGAGAGGCAAGACGCGGCGGAGTCTGCGAGACCAAGTGGAGG 60  
Db 164569 AGCGCCCTTTTGGTGGAGGGCAAGCGAGCGGGGAGTCTGGAGACCAAGTGGAGG 164510  
QY 61 AAGCGCGGAGATCTCAGGTTCCGCGCGGAGCCACTCTCTCTGTCGGCGGCTGCAGCG 120  
Db 164509 AAGCGCGGAGATCTCAGGTTCCGCGCGGAGCCACTCTCTCTGTCGGCGGCTGCAGCG 164450  
QY 121 GGGTGAAGCGGCGGAGCGGCGGAGTCTGAGACCATGGGCGGCGGCGGAGCGCATC 180  
Db 164449 GGGTGAAGCGGCGGAGCGGCGGAGTCTGAGACCATGGGCGGCGGCGGAGCGCATC 164390  
QY 181 CTGGATGCGCTGGAACACTGACCGCGGAGAGCTCAAGAAATTCAAGCTAAGCTGCTG 240  
Db 164389 CTGGATGCGCTGGAACACTGACCGCGGAGAGCTCAAGAAATTCAAGCTAAGCTGCTG 164330  
QY 241 TCGGTGCGCTGCGGAGGAGGCTACGGGCGCATCCGCGGCGGCGGCTGCTGTCATGAGC 300  
Db 164329 TCGGTGCGCTGCGGAGGAGGCTACGGGCGCATCCGCGGCGGCGGCTGCTGTCATGAGC 164270  
QY 301 GCCTTGAGACTTACCGCAACAGCTGTCTTACTTCTGGAACCTAAGAGCGCGGAGACTG 360  
Db 164269 GCCTTGAGACTTACCGCAACAGCTGTCTTACTTCTGGAACCTAAGAGCGCGGAGACTG 164210  
QY 361 ACCGCTAAGCTGCTGCGGAGCATGGGCTGAGAGATGGCGGCGAGCTGCAAGCGGCGC 420  
Db 164209 ACCGCTAAGCTGCTGCGGAGCATGGGCTGAGAGATGGCGGCGAGCTGCAAGCGGCGC 164150  
QY 421 ACGGACCAAGGATGAGCGCGCGCGCTTCCCTCCAGCGCGCTTTTCCCTCCAGCGCAAGC 480  
Db 164149 ACGGACCAAGGATGAGCGCGCGCGCTTCCCTCCAGCGCGCTTTTCCCTCCAGCGCAAGC 164090  
QY 481 AGCGCTTACCGCGCGGCGCTTCCCGCTTCTGTTCTCTTACCCCTAAACAAAGCTGCTC 540  
Db 164089 AGCGCTTACCGCGCGGCGCTTCCCGCTTCTGTTCTCTTACCCCTAAACAAAGCTGCTC 164030  
QY 541 TACCGGAAAGAGAGCTCCCGACAGCTTGGCTTACCGCAACAGGAGCGCGCGCGCAAGGC 600  
Db 164029 TACCGGAAAGAGAGCTCCCGACAGCTTGGCTTACCGCAACAGGAGCGCGCGCGCAAGGC 163970  
QY 601 GGGAGGGAAGGAGGAGGATCACTT 626  
Db 163969 GGGAGGGAAGGAGGAGGATCACTT 163944

RESULT 5  
LOCUS AX017270 779 bp DNA linear PAT 07-SEP-2000  
DEFINITION Sequence 21 from Patent WO9947669.  
ACCESSION AX017270

VERSION AX017270.1 GI:10042188

#### KEYWORDS

SOURCE human.

#### ORGANISM

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Primates; Catarrhini; Homidae; Homo.

REFERENCE 1 (bases 1 to 779)  
AUTHORS Schmitt,A., Specht,T., Dahl,E., Hinzmann,B., Rosenthal,A. and  
Pilarczyk,C.

TITLE Human nucleic acid sequences from tissue of breast tumors  
JOURNAL Patent: WO 9947669-A 21.23-SEP-1999.

SCHMITT ARMIN (DE); SPECHT THOMAS (DE); DAHL ENGAR (DE); HINZMANN  
BERND (DE); ROSENTHAL ANDRE (DE); METZGEN GES FUER GENOMFORSCHUN  
(DE); PILARSKY CHRISTIAN (DE)  
FEATURES  
source 1. 779  
/organism="Homo sapiens"  
/db\_xref="taxon:9606"

BASE COUNT 149 a 252 c 251 g 127 t  
ORIGIN

Query Match 58.8%; Score 368; DB 6; Length 779;  
Best Local Similarity 100.0%; Pred. No. 9.8e-41;  
Matches 368; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 64 GGGGAGATCCAGGTTCCGCGCGGAGCGGAGCTTCTCTGTCGGGCGGTGAGGCGG 123  
Db 1 GGGGAGATCCAGGTTCCGCGCGGAGCGGAGCTTCTCTGTCGGGCGGTGAGGCGG 60  
QY 124 TGAGGCGGAGCGGCGGCGGAGTCTGAGGACATGGGCGGCGGCGGAGCATCTG 183  
Db 61 TGAGGCGGAGCGGCGGCGGAGTCTGAGGACATGGGCGGCGGCGGAGCATCTG 120  
QY 184 GATGCGCTGAGAACTGACCGCGGAGAGCTCAAGAAATTCAAGCTGAAGCTGCTG 243  
Db 121 GATGCGCTGAGAACTGACCGCGGAGAGCTCAAGAAATTCAAGCTGAAGCTGCTG 180  
QY 244 GTGCGCTGCGGAGGCGTACGGGCGCATCCGCGGCGGCGGCTGCTGTCATGAGCGC 303  
Db 181 GTGCGCTGCGGAGGCGTACGGGCGCATCCGCGGCGGCGGCTGCTGTCATGAGCGC 240  
QY 304 TTGGAGCTACGAGCAAGCTGTGACCTTCTACCTGAGAGCTACAGGCGGAGCTAC 363  
Db 241 TTGGAGCTACGAGCAAGCTGTGACCTTCTACCTGAGAGCTACAGGCGGAGCTAC 300  
QY 364 GCTAAGCTGTCGCGGAGCATGGGCTGAGAGATGGCGGCGAGCTGCAAGCGGCGAGC 423  
Db 301 GCTAAGCTGTCGCGGAGCATGGGCTGAGAGATGGCGGCGAGCTGCAAGCGGCGAGC 360  
QY 424 CACGAGG 431  
Db 361 CACGAGG 368

RESULT 6  
LOCUS AX459863 782 bp DNA linear PAT 08-JUL-2002  
DEFINITION Sequence 6 from Patent WO0240668.

ACCESSION AX459863  
VERSION AX459863.1 GI:21725640  
KEYWORDS

SOURCE human.

#### ORGANISM

Homo sapiens  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Primates; Catarrhini; Homidae; Homo.

REFERENCE 1  
AUTHORS Tschopp,J. and Martinon,F.

TITLE Proteins and dna sequences underlying these proteins used for  
treating inflammations  
JOURNAL Patent: WO 0240668-A 6 23-MAY-2002;

Apotech Research and Development Ltd. (CH)  
FEATURES  
source 1. 782  
Location/Qualifiers

/organism="Homo sapiens"  
/db\_xref="taxon:9606"  
/note="Pycard.cdna"  
BASE COUNT 158 a 251 c 246 g 127 t  
ORIGIN

Query Match 57.7%; Score 361; DB 6; Length 782;  
Best Local Similarity 100.0%; Pred. No. 8.6e-40;  
Matches 361; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 71 GTCCAGTTCCGCCCGGAGCCGACCTTCTCTGCTGCGCGGCTGCGACGGGGTGAACGG 130  
DB 1 GTCCAGTTCCGCCCGGAGCCGACCTTCTCTGCTGCGCGGCTGCGACGGGGTGAACGG 60  
QY 131 CGGACGGCGCGGGGATCTGAGACCATGGGGCGCGCGCGGACGCCATCTGATGATCGG 190  
DB 61 CGGACGGCGCGGGGATCTGAGACCATGGGGCGCGCGCGGACGCCATCTGATGATCGG 120  
QY 191 TGGAGAACCTGACCGCCGAGAGCTCAAGATTCAAGCTGAAGCTGTGTGGTGGCGG 250  
DB 121 TGGAGAACCTGACCGCCGAGAGCTCAAGATTCAAGCTGAAGCTGTGTGGTGGCGG 180  
QY 251 TGGCGAGGGCTACGCGGCGCATCCCGCGGGCGCGCTGCTGTCAATGAGACCCCTTGAGC 310  
DB 181 TGGCGAGGGCTACGCGGCGCATCCCGCGGGCGCGCTGCTGTCAATGAGACCCCTTGAGC 240  
QY 311 TCACCGACAAGCTGTGACCTTCTACCTGAGACCTACAGCGCGCGGAGCTCACCGCTTAAG 370  
DB 241 TCACCGACAAGCTGTGACCTTCTACCTGAGACCTACAGCGCGCGGAGCTCACCGCTTAAG 300  
QY 371 TGTGTGGCGACATGGGCTCTGACAGAGATGGCCGGGAGCTGACAGGCGGCACGACCAAG 430  
DB 301 TGTGTGGCGACATGGGCTCTGACAGAGATGGCCGGGAGCTGACAGGCGGCACGACCAAG 360  
QY 431 G 431  
DB 361 G 361

RESULT 7 782 bp mRNA linear PRI 18-OCT-2000  
AB023416  
LOCUS Homo sapiens ASC mRNA for apoptosis-associated speck-like protein  
DEFINITION containing a CARD, complete cds.  
ACCESSION AB023416  
VERSION AB023416.2 GI:10801601  
KEYWORDS ASC.  
SOURCE Homo sapiens CDNA to mRNA.  
ORGANISM Homo sapiens  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Primates; Catarrhini; Hominiidae; Homo.  
REFERENCE 1 (sites)  
AUTHORS Masumoto,J., Taniguchi,S., Ayukawa,K., Sarvotham,H., Kishino,T.,  
TITLE Nilkawa,N., Hidaka,E., Katsuyama,T., Higuchi,T. and Sagara,J.  
JOURNAL ASC, a novel 22-kDa protein, aggregates during apoptosis of human  
MEDLINE promyelocytic leukemia HL-60 cells  
REFERENCE 2 (bases 1 to 782)  
AUTHORS Masumoto,J., Sagara,J. and Taniguchi,S.  
TITLE Direct Submission  
JOURNAL Submitted (04-FEB-1999) Junya Masumoto, Shinsu University School  
of Medicine, Research Center on Aging and Adaptation; Asahi 3-1-1,  
Matsumoto, Nagano 390-8621, Japan  
(E-mail:masumoto@sch.md.shinsu-u.ac.jp, Tel:81-263-37-2723,  
Fax:81-263-37-2724)  
COMMENT On Oct 14, 2000 this sequence version replaced gi:6482371.  
FEATURES  
SOURCE location/qualifiers  
1..782  
/organism="Homo sapiens"  
/db\_xref="taxon:9606"  
/chromosome="16"

/map="16p11.2-12"  
1..782  
/gene="ASC"  
87..674  
/gene="ASC"  
/codon\_start=1  
/product="apoptosis-associated speck-like protein  
containing a CARD"  
/protein\_id="BA087339.2"  
/db\_xref="GI:10801602"  
/translation="MGARDAITLDALENTLAEELKKFKLTVPLREGYRIPRGL  
LSMDALDITDKLVSPFLYETGAEITANNVMDGLOEMAGCOIQAATHGGAAPAGIOA  
PPQSAKPELHFIDHRAILIAKYVWEWLDALIGKVLTDBOYQVRAEPTPSMKR  
KLFSFTPMNMVTKDLLOALRESQSYLVEDLERS"

BASE COUNT 158 a 251 c 246 g 127 t  
ORIGIN

Query Match 57.7%; Score 361; DB 9; Length 782;  
Best Local Similarity 100.0%; Pred. No. 8.6e-40;  
Matches 361; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 71 GTCCAGTTCCGCCCGGAGCCGACCTTCTCTGCTGCGCGGCTGCGACGGGGTGAACGG 130  
DB 1 GTCCAGTTCCGCCCGGAGCCGACCTTCTCTGCTGCGCGGCTGCGACGGGGTGAACGG 60  
QY 131 CGGACGGCGCGGGGATCTGAGACCATGGGGCGCGCGCGGACGCCATCTGATGATCGG 190  
DB 61 CGGACGGCGCGGGGATCTGAGACCATGGGGCGCGCGCGGACGCCATCTGATGATCGG 120  
QY 191 TGGAGAACCTGACCGCCGAGAGCTCAAGATTCAAGCTGAAGCTGTGTGGTGGCGG 250  
DB 121 TGGAGAACCTGACCGCCGAGAGCTCAAGATTCAAGCTGAAGCTGTGTGGTGGCGG 180  
QY 251 TGGCGAGGGCTACGCGGCGCATCCCGCGGGCGCGCTGCTGTCAATGAGACCCCTTGAGC 310  
DB 181 TGGCGAGGGCTACGCGGCGCATCCCGCGGGCGCGCTGCTGTCAATGAGACCCCTTGAGC 240  
QY 311 TCACCGACAAGCTGTGACCTTCTACCTGAGACCTACAGCGCGCGGAGCTCACCGCTTAAG 370  
DB 241 TCACCGACAAGCTGTGACCTTCTACCTGAGACCTACAGCGCGCGGAGCTCACCGCTTAAG 300  
QY 371 TGTGTGGCGACATGGGCTCTGACAGAGATGGCCGGGAGCTGACAGGCGGCACGACCAAG 430  
DB 301 TGTGTGGCGACATGGGCTCTGACAGAGATGGCCGGGAGCTGACAGGCGGCACGACCAAG 360  
QY 431 G 431  
DB 361 G 361

RESULT 8 713 bp DNA linear PAT 11-MAY-2001  
AX118641  
LOCUS Sequence 24 from Patent WO0129235.  
DEFINITION AX118641  
ACCESSION AX118641  
VERSION AX118641.1 GI:1403592  
KEYWORDS  
SOURCE human.  
ORGANISM Homo sapiens  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Primates; Catarrhini; Hominiidae; Homo.  
REFERENCE 1 (bases 1 to 713)  
AUTHORS Vertino,F.W.  
TITLE Tms1 compositions and methods of use  
JOURNAL Patent: WO 0129235-A 24 26-APR-2001;  
Emory University (US)  
FEATURES  
SOURCE location/qualifiers  
1..713  
/organism="Homo sapiens"  
/db\_xref="taxon:9606"  
75..605  
/note="unnamed protein product"  
/codon\_start=1

```
/protein_id="CAC38476.1"
/db_xref="GI:14035593"
/translation="MGRARDAILDALLENLTAEIKRFLKLLSVLPREGYRIRPGAL
LSMDALDITDKLVSYFLTYGALNTANVLRDMQIQMAGOLQAAHQGLHFTIDHRA
LIARVTNVEMLDALYGVLTDEQYQAVRAEPTNPSPKMRKLFSTPANNMTCKDLLLO
ALRESOSYLVEDLERS"
BASE COUNT      148 a      224 c      223 g      118 t
ORIGIN
Query Match      54.8%; Score 342.8; DB 6; Length 713;
Best Local Similarity 98.0%; Pred. No. 2.5e-37;
Matches 347; Conservative 0; Mismatches 7; Indels 0; Gaps 0;
QY      83 CCGCGAGAGACCTCTCTCTGTGGTGGCGGCGTGCAGCGGGGTGAGCGCGGCGCGCG 142
      11 11 11 11 11 11 11 11 11 11 11 11 11 11 11 11 11 11 11 11 11 11
      1 CCACGGCTCCGACCTTCTCTCTGTGGTGGCGGCGTGCAGCGGGGTGAGCGCGGCGCGCG 60
QY      143 GGGATCTGAGACCATGGCGCGCGCGCGCGCATCTCTGATGCGCTGAGAACCTGA 202
      11 11 11 11 11 11 11 11 11 11 11 11 11 11 11 11 11 11 11 11 11 11
      61 GGGATCTGAGACCATGGCGCGCGCGCGCATCTCTGATGCGCTGAGAACCTGA 120
QY      203 CCGCGAGAGACCTCAAGAACTCAAGCTCTGTGCGTCCGCTCCGCGAGAGGCT 262
      11 11 11 11 11 11 11 11 11 11 11 11 11 11 11 11 11 11 11 11 11 11
      121 CCGCGAGAGACCTCAAGAACTCAAGCTCTGTGCGTCCGCTCCGCGAGAGGCT 180
QY      263 ACGGGCGCATCCCCGGGGCGCGCTCTGTCTGATGAGCGCTTGAGCTTACCGACAAGC 322
      11 11 11 11 11 11 11 11 11 11 11 11 11 11 11 11 11 11 11 11 11 11
      181 ACGGGCGCATCCCCGGGGCGCGCTCTGTCTGATGAGCGCTTGAGCTTACCGACAAGC 240
QY      323 TGGTACGCTTACCTGAGACCTAGCGCGCGAGCTACCGGCTACGCTGCGGAGCA 382
      11 11 11 11 11 11 11 11 11 11 11 11 11 11 11 11 11 11 11 11 11 11
      241 TGGTACGCTTACCTGAGACCTAGCGCGCGAGCTACCGGCTACGCTGCGGAGCA 300
QY      383 TGGGCGCTGAGAGATGGCGGCGAGCTGACAGCGCGCGACGACAGGCTGAC 436
      11 11 11 11 11 11 11 11 11 11 11 11 11 11 11 11 11 11 11 11 11 11
      301 TGGGCGCTGAGAGATGGCGGCGAGCTGACAGCGCGCGACGACAGGCTGAC 354
DB
RESULT 9
AF255794      713 bp      mRNA      11linear      PRI 02-FEB-2001
LOCUS
DEFINITION
Homo sapiens target of methylation-induced silencing-1 (TMS1) mRNA,
complete cds, alternatively spliced.
ACCESSION
AF255794.1      GI:9802275
VERSION
AF255794.1      GI:9802275
KEYWORDS
SOURCE
ORGANISM
Homo sapiens.
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
1 (bases 1 to 713)
REFERENCE
CONWAY,K.E., MCCOMMELL,B.B., BOWLING,C.E., DONALD,C.D., WARREN,S.T.
and VERTINO,P.M.
TMS1, a novel proapoptotic caspase recruitment domain protein, is a
target of methylation-induced gene silencing in human breast
cancers
JOURNAL
MEDLINE
20552139
PUBMED
11103776
AUTHORS
MCCOMMELL,B.B. and VERTINO,P.M.
TITLE
Activation of a caspase-9-mediated apoptotic pathway by subcellular
redistribution of the novel caspase recruitment domain protein TMS1
JOURNAL
Cancer Res. 60 (22), 6243-6247 (2000)
MEDLINE
20552140
PUBMED
11103777
AUTHORS
VERTINO,P.M.
TITLE
Direct Submission
JOURNAL
Submitted (13-APR-2000) Radiation Oncology, Emory University School
of Medicine, 145 Edgewood Avenue, SE, Atlanta, GA 30335, USA
FEATURES
source
1. .713
```

```
/organism="Homo sapiens"
/db_xref="taxon:9606"
/chromosome="16"
/map="16p11.2-p12.1"
/cell_line="Hs578st"
/tissue_type="normal breast epithelium"
1. .713
/gene="TMS1"
/gene="TMS1"
75. .605
/gene="TMS1"
/function="Involved in apoptosis"
/note="contains caspase-recruitment domain; CARD protein;
alternatively spliced form lacking exon 2; target of
methylation-mediated gene silencing in human tumors"
/codon_start=1
/product="target of methylation-induced silencing-1"
/protein_id="AF9655.1"
/db_xref="GI:9802276"
/translation="MGRARDAILDALLENLTAEIKRFLKLLSVLPREGYRIRPGAL
LSMDALDITDKLVSYFLTYGALNTANVLRDMQIQMAGOLQAAHQGLHFTIDHRA
LIARVTNVEMLDALYGVLTDEQYQAVRAEPTNPSPKMRKLFSTPANNMTCKDLLLO
ALRESOSYLVEDLERS"
682. .687
polyA_signal
BASE COUNT      148 a      224 c      223 g      118 t
ORIGIN
Query Match      54.8%; Score 342.8; DB 9; Length 713;
Best Local Similarity 98.0%; Pred. No. 2.5e-37;
Matches 347; Conservative 0; Mismatches 7; Indels 0; Gaps 0;
QY      83 CCGCGAGAGACCTCTCTCTGTGGTGGCGGCGTGCAGCGGGGTGAGCGCGGCGCGCG 142
      11 11 11 11 11 11 11 11 11 11 11 11 11 11 11 11 11 11 11 11 11 11
      1 CCACGGCTCCGACCTTCTCTCTGTGGTGGCGGCGTGCAGCGGGGTGAGCGCGGCGCGCG 60
QY      143 GGGATCTGAGACCATGGGGCGCGCGCGCGCATCTCTGATGCGCTGAGAACCTGA 202
      11 11 11 11 11 11 11 11 11 11 11 11 11 11 11 11 11 11 11 11 11 11
      61 GGGATCTGAGACCATGGGGCGCGCGCGCGCATCTCTGATGCGCTGAGAACCTGA 120
QY      203 CCGCGAGAGACCTCAAGAACTCAAGCTCTGTGCGTCCGCTCCGCGAGAGGCT 262
      11 11 11 11 11 11 11 11 11 11 11 11 11 11 11 11 11 11 11 11 11 11
      121 CCGCGAGAGACCTCAAGAACTCAAGCTCTGTGCGTCCGCTCCGCGAGAGGCT 180
QY      263 ACGGGCGCATCCCCGGGGCGCGCTCTGTCTGATGAGCGCTTGAGCTTACCGACAAGC 322
      11 11 11 11 11 11 11 11 11 11 11 11 11 11 11 11 11 11 11 11 11 11
      181 ACGGGCGCATCCCCGGGGCGCGCTCTGTCTGATGAGCGCTTGAGCTTACCGACAAGC 240
QY      323 TGGTACGCTTACCTGAGACCTAGCGCGCGAGCTACCGGCTACGCTGCGGAGCA 382
      11 11 11 11 11 11 11 11 11 11 11 11 11 11 11 11 11 11 11 11 11 11
      241 TGGTACGCTTACCTGAGACCTAGCGCGCGAGCTACCGGCTACGCTGCGGAGCA 300
QY      383 TGGGCGCTGAGAGATGGCGGCGAGCTGACAGCGCGCGACGACAGGCTGAC 436
      11 11 11 11 11 11 11 11 11 11 11 11 11 11 11 11 11 11 11 11 11 11
      301 TGGGCGCTGAGAGATGGCGGCGAGCTGACAGCGCGCGACGACAGGCTGAC 354
DB
RESULT 10
AX118643      405 bp      DNA      11linear      PAT 11-MAY-2001
LOCUS
DEFINITION
Sequence 26 from Patent WO0129235.
ACCESSION
AX118643
VERSION
AX118643.1      GI:14035594
KEYWORDS
SOURCE
human.
ORGANISM
Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
1 (bases 1 to 405)
REFERENCE
VERTINO,P.M.
TMS1 compositions and methods of use
Patent: WO 0129235-A 26 26-APR-2001;
Emory University (US)
FEATURES
Location/Qualifiers
```

**BASE COUNT**

gene

```
/gene="TMS1"
```

	CDS	75..662 /gene="TMS1" /function="Involved in apoptosis" /note="contains caspase-recruitment domain; CARD protein; target of methylation-mediated gene silencing in Homo sapiens tumors" /codon-start=1 /product="target of methylation-induced silencing 1" /protein_id="AAG01188.1" /db_xref="GI:9863864"
	BASE COUNT	157 a 247 c 241 g 125 t
	ORIGIN	
Query Match	54.7%;	Score 342.6; DB 9;
Best Local Similarity	98.9%;	Pred. No. 2.6e-37;
Matches 345;	Conservative 0;	Mismatches 4; Indels 0; Gaps 0;
OY	83 CCCCCGACCCACTTCCTCCTCGTGCGGGGTGCACGCGGGGTAGCGCGGCGGCCG 142	
Db	1 CCACGGCGTCCGACTTCCTCGTGCGGGGTGCACGCGGGGTAGCGCGGCGGCCG 60	
OY	143 GGATCATCGAACCATGGGCGCGCGCGCACGCCATCTCGAAATGGCTGTGGGAACCTGA 202	
Db	61 GGATCATCGAACCAATTGGGGCGCGCGCGCACGCCATCTCGAATGGCTGTGGGAACCTGA 120	
OY	203 CCCCAGAGAGCTCAAGAAGTTTAAGCTTAGCTGTGTGGTCCGCTCGCGAGAGGCT 262	
Db	121 CGCCCGAGAGAGCTCAAGAAGTTTAAGCTTAGCTGTGTGGTCCGCTCGCGAGAGGCT 180	
OY	263 ACGGGCGCATCCCGCGGGGCGGCGCTCTGTCATGAGAACGCCCTTGAGACTCAACGACAAG 322	
Db	181 ACGGGCGCATCCCGGGGCGGCGCTCTGTCATGAGAACGCCCTTGAGACTCAACGACAAG 240	
OY	323 TGGTCAGCTTACTACTGGAGACTTACGGCGCGAGACTACCAGTAACTGTGTGGGCACA 382	
Db	241 TGTCAGCTTACTACTGGAGACTTACGGGGCGAGACTACCAGTAACTGTGTGGGCACA 300	
OY	383 TGGGCGCTGCAGAGAGATGGCCGGGCACTGTCAGAGGGCGCCACGACCAAGG 431	
Db	301 TTGGCGCTGCAGAGAGATGGCCGGGCACTGTCAGAGGGCGCCACGACCAAGG 349	
RESULT 13		
AXI18622		
LOCUS	AXI18622	340 bp DNA linear PAT 11-MAY-2001
DEFINITION	Sequence 5 from Patent WO0129235.	
ACCESSION	AXI18622	
VERSION	AXI18622.1	GI:14035573
KEYWORDS	human. human. Homo sapiens	
ORGANISM	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominiidae; Homo.	
REFERENCE	1 (bases 1 to 340)	
AUTHORS	Vertino,P.M.	
TITLE	Tmsl compositions and methods of use	
JOURNAL	Patent: WO 0129235-A 5 26-Apr-2001;	
FEATURES	Emory University (US)	
source	location/Qualifiers	
	1..340	
	/organism="Homo sapiens"	
	/db_xref="taxon:9606"	
	67..>339	
	/note="unnamed protein product"	
	/codon_start=1	
	/protein_id="CAC38472.1"	
	/db_xref="GI:14035574"	
	/translation="MGARAADIIDALENLTAREBKRRKLLSVPLREGYGIRPGAL LSMDADLTDKLVSFYLETYLETGYAEITAVNLRMGLQENAGOLQAATHO"	

BASE COUNT	53 a	112 c	127 g	48 t
ORIGIN				
Query Match	54.3%	Score 340;	DB 6;	Length 740;
Best Local Similarity	100.0%;	Pred. No. 7.2e-37;		
Matches 340;	Conservative 0;	Mismatches 0;	Indels 0;	Gaps 0;
OY	91	CCGACTTCTCTCTGTGTGGGGGCTGCAGAGGGGGTGTAGCGCGACGGCCGGGGATCTT	150	
DB	1	CCCACTTCTCTCTGTGTGGGGGCTGCAGAGGGGGTGTAGCGCGACGGCGGATCTT	60	
OY	151	GGAGCCATGGGGGGGGGGGGGGGAGCGGCATCTCGGATGCGTGGAGAAAGCTGACGGCCGAG	210	
DB	61	GGAGCCATGGGGGGGGGGGGGGGAGCGGCATCTCGGATGCGTGGAGAAAGCTGACGGCCGAG	120	
OY	211	GAGCTCAGAGAGTTTCAGCTGAAGCTGCTGTGTGTGTCGCGCTGCGCGAGGGCTACGGGGCG	270	
DB	121	GAGCTCAGAGAGTTTCAGCTGAAGCTGCTGTGTGTGTCGCGCGCTGCGCGAGGGCTACGGGGCG	180	
OY	271	ATCCGGCGGG	330	
DB	181	ATCCGGCGGG	240	
OY	331	TTCCTACCTGAGACCTACGCGCGCGGACGCTCAACGCTGCTGCGCGACATGGGGCTG	390	
DB	241	TTCCTACCTGAGACCTACGCGCGCGGACGCTCAACGCTGCTGCGCGACATGGGGCTG	300	
OY	391	CAGGAGATGGCGGGGCGGAGCTGCAGAGGGGGCGACGACGACG	450	
DB	301	CAGGAGATGGCGGGGCGGAGCTGCAGAGGGGGCGACGACGACG	340	
RESULT 14				
LOCUS	AX082246	740 bp	DNA	linear
DEFINITION	Sequence 48 from Patent WO0100826.			
ACCESSION	AX082246			
VERSION	AX082246.1	GI:13171007		
KEYWORDS	human.			
SOURCE	human.			
ORGANISM	Homo sapiens			
REFERENCE	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;			
AUTHORS	Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.			
TITLE	Bertin,J.			
JOURNAL	Novel molecules of the card-related protein family and uses thereof			
FEATURES	Patent: NO 0100826-A 48 04-JAN-2001;			
source	Millennium Pharmaceuticals, Inc. (US)			
	Location/Qualifiers			
	1..740			
	/organism="Homo sapiens"			
	/db_xref="taxon:9606"			
	54..641			
CDS				
	/note="unnamed protein product"			
	/codon_start=1			
	/protein_id="CAC33147.1"			
	/db_xref="GI:13171008"			
	/translation="MGRARDAIIDLLENLTAEELKFKELKLSVPLREGYRIPRGAL			
	LSMDALDLTKIVSYLETYGAIDITANVLRDMGLQEMAGOLQATFHOGGAAPAGIOA			
	PPQSAKPELHFIIDQHRALIAIVTVEMLIDALYGVLTDEQYQAVRAPEPTNSKMR			
	KLESPFPMNMTCDLILQALRESQYLVDELNS"			
BASE COUNT	146 a	238 c	236 g	120 t
ORIGIN				
Query Match	51.4%;	Score 322;	DB 6;	Length 740;
Best Local Similarity	100.0%;	Pred. No. 1.6e-34;		
Matches 322;	Conservative 0;	Mismatches 0;	Indels 0;	Gaps 0;
OY	110	CGGCTGCAGCGGGGTGAGCGGGGCGGAGCGGGCGGATCTCTGAGACCAATGGGGCGGGCG	169	
DB	7	CGGCTGCAGCGGGGTGAGCGGGGCGGAGCGGGCGGATCTCTGAGACCAATGGGGCGGGCG	66	
OY	170	GGGAGCGCCATCTTGATGCGCTGGAAGCACTGACCGCGGAGGAGCTCAAGAAATTCAAG	229	



Db 67 GCAGCCATCTCGATGCGTGGAGAACCTGACCGCGAGGACTCAAAAGTTCAAGC 126  
QY 230 TGAAGCTGCTGTGCGTGCCGCTGCGCGAGGCGCTACGGGCGCATCCCGGGGCGCGCTGC 289  
Db 127 TGAAGCTGCTGTGCGTGCCGCTGCGCGAGGCGCTACGGGCGCATCCCGGGGCGCGCTGC 186  
QY 290 TGTCCATGAGAGCGCTTGAGCTCAACGACACGCTGACGCTTACCTGAGAGACCTAGC 349  
Db 187 TGTCCATGAGAGCGCTTGAGCTCAACGACACGCTGAGCTTACCTGAGAGACCTAGC 246  
QY 350 GCGCCGAGCTACCGCTTAACGCTGCGCGAGACATGGGCGCTCAGAGAGATGCGGCGCAGC 409  
Db 247 GCGCCGAGCTACCGCTTAACGCTGCGCGAGACATGGGCGCTCAGAGAGATGCGGCGCAGC 306  
QY 410 TGCAGGCGGCGCACGACAGG 431  
Db 307 TGCAGGCGGCGCACGACAGG 328

RESULT 15  
AF384665 740 bp mRNA linear PRI 19-JUN-2001  
LOCUS Homo sapiens caspase recruitment domain protein 5 mRNA, complete  
DEFINITION cds.

ACCESSION AF384665  
VERSION AF384665.1 GI:14488058

KEYWORDS  
SOURCE Homo sapiens.

ORGANISM Homo sapiens  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.

REFERENCE  
TITLE CARD5 protein is a CARD/PYRIN family member that is involved in  
apoptosis signal transduction

JOURNAL  
REFERENCE  
TITLE CARD5 protein is a CARD/PYRIN family member that is involved in  
apoptosis signal transduction

AUTHORS  
TITLE Direct Submission  
JOURNAL Submitted (23-MAY-2001) Neurobiology, Millennium Pharmaceuticals  
Inc., 640 Memorial Drive, Cambridge, MA 02139, USA

FEATURES  
source  
1..740  
Location/Qualifiers

CDS  
/organism="Homo sapiens"  
/db\_xref="taxon:9606"  
54..641  
/note="CARD5; CARD/PYRIN family member; bipartite protein  
comprised of an N-terminal PYRIN domain and a C-terminal  
CARD domain; involved in apoptosis signal transduction"

/codon\_start=1  
/product="caspase recruitment domain protein 5"  
/protein\_id="AAK63850.1"  
/db\_xref="GI:14488058"  
/translation="MGKARDATLDALENTAEIKKKKLLSVPLRBYGRIRPGAL  
LSMDALDTKIVSYLFETYGAEITANVLRMGLOEMAGQLAATHQSGAAPAGIOA  
PPOSAKRGHIFIDQHRALILARVNVEMLDALYGVITDQYQAVRAEPTNSKMR  
KLFSTPAMWMTCKDLICALLRESQSYLVEDLERS"

BASE COUNT 146 a 238 c 236 g 120 t

Query Match 51.4% Score 322; DB 9; Length 740;  
Best Local Similarity 100.0%; Pred. No. 1.e-34;  
Matches 322; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 110 CCGCTGACAGCGGGTGAAGCGGCGGCGGCGGATCTGAGCCATGGGGCGCGCGC 169  
Db 7 CCGCTGACAGCGGGTGAAGCGGCGGCGGCGGATCTGAGCCATGGGGCGCGCGC 66  
QY 170 GCGAGCGCACTGCTGAGAGAGCTGACCGCGGAGAGCTCAAGAGTTCAAGC 229  
Db 67 GCGAGCGCACTGCTGAGAGAGCTGACCGCGGAGAGCTCAAGAGTTCAAGC 126

QY 230 TGAAGCTGCTGTGCGTGCCGCTGCGCGAGGCGCTACGGGCGCATCCCGGGGCGCGCTGC 289  
Db 127 TGAAGCTGCTGTGCGTGCCGCTGCGCGAGGCGCTACGGGCGCATCCCGGGGCGCGCTGC 186  
QY 290 TGTCCATGAGAGCGCTTGAGCTCAACGACACGCTGACGCTTACCTGAGAGACCTAGC 349  
Db 187 TGTCCATGAGAGCGCTTGAGCTCAACGACACGCTGAGCTTACCTGAGAGACCTAGC 246  
QY 350 GCGCCGAGCTACCGCTTAACGCTGCGCGAGACATGGGCGCTCAGAGAGATGCGGCGCAGC 409  
Db 247 GCGCCGAGCTACCGCTTAACGCTGCGCGAGACATGGGCGCTCAGAGAGATGCGGCGCAGC 306  
QY 410 TGCAGGCGGCGCACGACAGG 431  
Db 307 TGCAGGCGGCGCACGACAGG 328

Search completed: December 22, 2002, 06:13:34  
Job time : 2750 secs

...

\_\_\_\_\_

\_\_\_\_\_

GenCore version 5.1.3  
Copyright (c) 1993 - 2002 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: December 22, 2002, 05:21:01 | Search time 257 Seconds  
(without alignments)  
5485.412 Million cell updates/sec

Title: US-09-691-763B-4

Perfect score: 626  
Sequence: 1 agcgccctttgcgtgagggc.....ggaaagggaaggatcatt 626

Scoring table: IDENTITY\_NUC  
Gapop 10.0, Gapext 1.0

Searched: 2185239 seqs, 112599159 residues

Total number of hits satisfying chosen parameters: 4370478

Minimum DB seq length: 0  
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%  
Maximum Match 100%

Listing first 45 summaries

Database:

N\_Geneseq\_101002: \*  
1: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1980.DAT: \*  
2: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1981.DAT: \*  
3: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1982.DAT: \*  
4: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1983.DAT: \*  
5: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1984.DAT: \*  
6: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1985.DAT: \*  
7: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1986.DAT: \*  
8: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1987.DAT: \*  
9: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1988.DAT: \*  
10: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1989.DAT: \*  
11: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1990.DAT: \*  
12: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1991.DAT: \*  
13: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1992.DAT: \*  
14: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1993.DAT: \*  
15: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1994.DAT: \*  
16: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1995.DAT: \*  
17: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1996.DAT: \*  
18: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1997.DAT: \*  
19: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1998.DAT: \*  
20: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1999.DAT: \*  
21: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA2000.DAT: \*  
22: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA2001A.DAT: \*  
23: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA2001B.DAT: \*  
24: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA2002.DAT: \*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match Length	ID	Description
1	626	100.0	626	22 AAD03891 Cpg island of huma
2	626	100.0	2821	22 AAD03889 Human target of me
3	389	62.1	806	22 AAH34052 Human colon cancer
4	389	62.1	811	21 AAC77884 Human cancer assoc
5	368	58.8	779	20 AA33631 Human breast tumor
6	361	57.7	782	24 AA47126 Pylori domain conta
7	342.8	54.8	713	22 AAD03906 Alternatively splic
8	342.6	54.7	405	22 AAD03907 Human target of me
9	342.6	54.7	770	22 AAD03890 Human target of me

10	340	54.3	340	22 AAD03892 Human target of me
11	322	51.4	740	22 AAF30007 Human CARD-5 CDNA.
12	322	51.4	740	24 ABK87966 Human caspase recr
13	322	51.4	740	24 ABK87967 Human caspase recr
14	321	51.3	745	22 AAS59817 Human novel cytoki
15	320	51.1	432	21 AAC98638 Human colon cancer
16	178.8	28.6	777	21 AAC98638 Human CARD-5 CDNA.
17	178.8	28.6	777	24 ABK87964 Mouse caspase recr
18	178.8	28.6	777	24 ABK87965 Mouse caspase recr
19	178.6	28.5	803	22 AAD03904 Mouse target of me
20	177.6	28.4	551	24 AA47124 Pylori domain conta
21	174.4	27.9	807	22 AAC88099 Human FLEXHF-30 nu
22	174.4	27.9	807	22 AAC88099 Human colon cancer
23	174.4	27.9	807	22 AAC88099 Human colon cancer
24	174.4	27.9	807	22 AAC88099 Human colon cancer
25	174.4	27.9	807	22 AAC88099 Human colon cancer
26	174.4	27.9	807	22 AAC88099 Human colon cancer
27	174.4	27.9	807	22 AAC88099 Human colon cancer
28	174.4	27.9	807	22 AAC88099 Human colon cancer
29	174.4	27.9	807	22 AAC88099 Human colon cancer
30	174.4	27.9	807	22 AAC88099 Human colon cancer
31	174.4	27.9	807	22 AAC88099 Human colon cancer
32	174.4	27.9	807	22 AAC88099 Human colon cancer
33	174.4	27.9	807	22 AAC88099 Human colon cancer
34	174.4	27.9	807	22 AAC88099 Human colon cancer
35	174.4	27.9	807	22 AAC88099 Human colon cancer
36	174.4	27.9	807	22 AAC88099 Human colon cancer
37	174.4	27.9	807	22 AAC88099 Human colon cancer
38	174.4	27.9	807	22 AAC88099 Human colon cancer
39	174.4	27.9	807	22 AAC88099 Human colon cancer
40	174.4	27.9	807	22 AAC88099 Human colon cancer
41	174.4	27.9	807	22 AAC88099 Human colon cancer
42	174.4	27.9	807	22 AAC88099 Human colon cancer
43	174.4	27.9	807	22 AAC88099 Human colon cancer
44	174.4	27.9	807	22 AAC88099 Human colon cancer
45	174.4	27.9	807	22 AAC88099 Human colon cancer

ALIGNMENTS

RESULT 1	
AAD03891	
ID	AAD03891 standard; DNM: 626 BP.
AC	AAD03891;
DT	02-JUL-2001 (first entry)
DE	Cpg island of human TMS1 gene.
XX	
KW	Human: target of methylation-induced silencing-1; TMS1; cytostatic;
KW	antiproliferative; apoptosis inducer; gene therapy; Cpg island;
KW	caspase-recruiting domain; CARD; cancer; breast; ds.
XX	
OS	Homo sapiens.
PN	MO200129235-A2.
PD	26-APR-2001.
PF	18-OCT-2000; 2000MO-US28747.
PR	18-OCT-1999; 99US-0159975.
PA	(UYEM-) UNIV EMORY.
XX	
PI	Vertino PM;
XX	
DR	WPT; 2001-290922/30.
XX	
PT	Novel gene TMS1, transcriptionally silenced due to increased methylation useful for identifying subject at risk of developing tumor

PT characterized by abnormal methylation, for treating cancer by inducing  
PT apoptosis -

PS Claim 71: Page 114-115; 124pp; English.

CC The invention relates to identification of target of methylation-induced  
CC silencing-1 (TMS1) gene. This gene is transcriptionally silenced due to  
CC abnormal methylation of a CpG island in its 5' regulatory region. TMS1  
CC consists of a carboxy terminal caspase-recruiting domain (CARD) and  
CC plays a role in induction of apoptosis. TMS1 gene and protein are useful  
CC as tools for diagnosing and treating a subject at risk of developing  
CC cancer (e.g. breast cancer) characterised by abnormal CpG methylation or  
CC abnormally low levels of TMS1 expression products. Unique fragments of  
CC TMS1 gene are used as probes. TMS1 gene is useful in gene therapy.  
CC TMS1 molecule is also useful for treating abnormal cell proliferation by  
CC increasing TMS1 polypeptide level to an above normal level. The CpG  
CC island region of TMS1 or its fragments are used to study the methylation  
CC patterns apart from any coding region contained in it.  
CC The present sequence is CpG island of target of methylation-induced  
CC silencing-1 (TMS1) gene.

SQ Sequence 626 BP; 103 A; 216 C; 216 G; 91 T; 0 other;

Query Match 100.0%; Score 626; DB 22; Length 626;

Best Local Similarity 100.0%; Pred. No. 1,1e-103; Mismatches 0; Gaps 0;

Matches 626; Conservative 0; Indels 0; Gaps 0;

OY 1 AAGCGCTTTTCTGAGAGGCAACGAGCGGGGCGGGAGTCGGAGACCAAGATGGAGG 60  
DB 1 AAGCGCTTTTCTGAGAGGCAACGAGCGGGGCGGGAGTCGGAGACCAAGATGGAGG 60  
OY 61 AAGCGGGGAGTCGAGGTCGCGCGCGGAGCGACTTCCTGTCGGGCGGCTGCAGCG 120  
DB 61 AAGCGGGGAGTCGAGGTCGCGCGCGGAGCGACTTCCTGTCGGGCGGCTGCAGCG 120  
OY 121 GGGTGAAGCGGCGAGCGCGCGGGGATCTCTGAGCCATGGGGCGCGCGACCCATC 180  
DB 121 GGGTGAAGCGGCGAGCGCGCGGGGATCTCTGAGCCATGGGGCGCGCGACCCATC 180  
OY 181 CTGGATGCGCTGGAGAAACCTGACCGCGAGAGCTCAAGAAAGTTCAAGGCTGAAGCTG 240  
DB 181 CTGGATGCGCTGGAGAAACCTGACCGCGAGAGCTCAAGAAAGTTCAAGGCTGAAGCTG 240  
OY 241 TCGGTGCGCGTGGCGAGAGGCTACGAGGCGCATCCCGCGGGCGCGCTGTCATGAGAC 300  
DB 241 TCGGTGCGCGTGGCGAGAGGCTACGAGGCGCATCCCGCGGGCGCGCTGTCATGAGAC 300  
OY 301 GCGTTGAGCGTCAACGCAAGGCGGTCACTTACCTGAGAGACTACGAGCGCGAGGTC 360  
DB 301 GCGTTGAGCGTCAACGCAAGGCGGTCACTTACCTGAGAGACTACGAGCGCGAGGTC 360  
OY 361 ACCGCTAACGCTGCGAGCAATGGGCTGCGAGAGATGGCCGCGACGTCGAGCGGCC 420  
DB 361 ACCGCTAACGCTGCGAGCAATGGGCTGCGAGAGATGGCCGCGACGTCGAGCGGCC 420  
OY 421 AAGCGACGAGGAGCGCGCGCGCTTCCCTTCCCTTCCCTTCCCTTCCCTTCCCTTCC 480  
DB 421 AAGCGACGAGGAGCGCGCGCGCTTCCCTTCCCTTCCCTTCCCTTCCCTTCCCTTCC 480  
OY 481 AAGCGCTTACCCCGCGGCTTCCCTTCTGTCCTTCTTCTTCTTCTTCTTCTTCTTCT 540  
DB 481 AAGCGCTTACCCCGCGGCTTCCCTTCTGTCCTTCTTCTTCTTCTTCTTCTTCTTCT 540  
OY 541 TACCGGAAAGAGAGCTCCACGCTTGGCTTACGACCAAGGAGCGCGCGCGCGCGCG 600  
DB 541 TACCGGAAAGAGAGCTCCACGCTTGGCTTGGCTTACGACCAAGGAGCGCGCGCGCG 600  
OY 601 GGGAGGGAAGGAGGAGGATCATCTT 626  
DB 601 GGGAGGGAAGGAGGAGGATCATCTT 626

RESULT 2

AAD03889  
ID AAD03889 standard; DNA; 2621 BP.  
XX  
AC AAD03889;  
XX  
PT 02-JUL-2001 (first entry)  
XX  
DE Human target of methylation-induced silencing-1 (TMS1) genomic DNA.  
XX  
KW Human; target of methylation-induced silencing-1; TMS1; cytostatic;  
KW antiproliferative; apoptosis inducer; gene therapy; CpG island;  
XX caspase-recruiting domain; CARD; cancer; breast; ds.  
OS Homo sapiens.  
XX  
FH Key location/qualifiers  
FH misc\_feature 1100..1725  
FT  
FT /tag= a  
FT /label= CpG\_island  
FT /note= "This region is specifically claimed in claim 71"  
FT CDS 1256..2361  
FT  
FT /tag= b  
FT /product= "Human TMS1 protein"  
FT /note= "CDS is interrupted by 2 introns"  
FT exon 1256..1529  
FT  
FT /tag= c  
FT /number= 1  
FT intron 1530..1742  
FT  
FT /tag= d  
FT /number= 1  
FT exon 1743..1799  
FT  
FT /tag= e  
FT /number= 2  
FT intron 1800..2104  
FT  
FT /tag= f  
FT /number= 2  
FT exon 2105..2361  
FT  
FT /tag= g  
FT /number= 3  
XX  
PN W0200129235-A2.  
XX  
PD 26-APR-2001.  
XX  
PE 18-OCT-2000; 2000WO-US28747.  
XX  
PR 18-OCT-1999; 99US-0159975.  
XX  
PA (UYEM-) UNIV EMORY.  
XX  
PI Vertino PM;  
XX  
XX WPI: 2001-290922/30.  
XX P-PSDB: AAB00588.  
XX  
PT Novel gene TMS1, transcriptionally silenced due to increased  
PT methylation useful for identifying subject at risk of developing tumor  
PT characterized by abnormal methylation, for treating cancer by inducing  
PT apoptosis -  
XX  
PS Claim 69: Page 112-113; 124pp; English.  
XX  
XX  
XX The invention relates to identification of target of methylation-induced  
XX silencing-1 (TMS1) gene. This gene is transcriptionally silenced due to  
XX abnormal methylation of a CpG island in its 5' regulatory region. TMS1  
XX consists of a carboxy terminal caspase-recruiting domain (CARD) and  
XX plays a role in induction of apoptosis. TMS1 gene and protein are useful  
XX as tools for diagnosing and treating a subject at risk of developing  
XX cancer (e.g. breast cancer) characterised by abnormal CpG methylation or  
XX abnormally low levels of TMS1 expression products. Unique fragments of  
XX TMS1 gene are used as probes. TMS1 gene is useful in gene therapy.  
XX TMS1 molecule is also useful for treating abnormal cell proliferation by  
XX increasing TMS1 polypeptide level to an above normal level. The CpG

CC Island region of TMS1 or its fragments are used to study the methylation  
CC patterns apart from any coding region contained in it.  
CC The present sequence is human target of methylation-induced silencing-1  
CC (TMS1) genomic DNA.

XX Sequence 2821 BP; 614 A; 864 C; 757 G; 586 T; 0 other;

Query Match 100.0%; Score 626; DB 22; Length 2821;  
Best Local Similarity 100.0%; Pred. No. 1.1e-103;  
Matches 626; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 AGCCCTTTTCTGAGGAGCAACGACGCGGGGAGTGGGAGACCAAGTGGAGG 60  
DB 1100 AGCGCTTTTCTGAGGAGCAACGACGCGGGGAGTGGGAGACCAAGTGGAGG 1159  
QY 61 AAGCGGGAGTCCAGTTCCGCCCGGAGCGCACTTCTCTGTCGGCGCTGCAGCG 120  
DB 1160 AAGCGGGAGTCCAGTTCCGCCCGGAGCGCACTTCTCTGTCGGCGCTGCAGCG 1219  
QY 121 GGGTAGCGGGGAGCGGCGGGGATCCTGAGACCATGGGCGCGCGGACGCAATC 180  
DB 1220 GGGTAGCGGGGAGCGGCGGGGATCCTGAGACCATGGGCGCGCGGACGCAATC 1279  
QY 181 CTGATGCGCTGAGAGACCTGACCGCGGAGAGCTCAAGATTCAAGCTGCTG 240  
DB 1280 CTGATGCGCTGAGAGACCTGACCGCGGAGAGCTCAAGATTCAAGCTGCTG 1339  
QY 241 TGGGTCCCGCTGCGGCGGAGGCTACGGGCGCATCCCGGGGGCGCGCTGTCATGAC 300  
DB 1340 TGGGTCCCGCTGCGGCGGAGGCTACGGGCGCATCCCGGGGGCGCGCTGTCATGAC 1399  
QY 301 GCGTTGACCTCACCGACAGCTGTCTACCTTCTACCTGAGACCTACGGCGCGAGCTC 360  
DB 1400 GCGTTGACCTCACCGACAGCTGTCTACCTTCTACCTGAGACCTACGGCGCGAGCTC 1459  
QY 361 ACCGCTTACGTCGCGGCGGACATGGGCTGAGAGAGTGGCGGAGCTGAGAGGCGGC 420  
DB 1460 ACCGCTTACGTCGCGGCGGACATGGGCTGAGAGAGTGGCGGAGCTGAGAGGCGGC 1519  
QY 421 ACGCACGAGGCTGAGCGCGCGCGCTCCCTCCACCGCGCTCTTCCCTCCACCGCAC 480  
DB 1520 ACGCACGAGGCTGAGCGCGCGCGCTCCCTCCACCGCGCTCTTCCCTCCACCGCAC 1579  
QY 481 ACGGTTTACCGCGGCGGCTCTTCCGTTCTGTTCTCTTCCCTTAAACAAAGCTGCTC 540  
DB 1580 ACGGTTTACCGCGGCGGCTCTTCCGTTCTGTTCTCTTCCCTTAAACAAAGCTGCTC 1639  
QY 541 TACCGGAAAGAGGCTCCCAAGCTTGGCTTACCGACCAACGAGGAGCGCGCGCGACG 600  
DB 1640 TACCGGAAAGAGGCTCCCAAGCTTGGCTTACCGACCAACGAGGAGCGCGCGCGACG 1699  
QY 601 GGGAGGGAAGGGAAGGAGTCACTT 626  
DB 1700 GGGAGGGAAGGGAAGGAGTCACTT 1725

RESULT 3  
AAH34052  
ID AAH34052 standard; cdna; 806 BP.

XX AAH34052;

XX 03-SEP-2001 (first entry)

XX Human colon cancer antigen encoding cdna seq ID NO:1134.

XX Human: colon cancer; colon cancer antigen; diagnosis; detection;

XX colorectal carcinoma; ss.

OS Homo sapiens.

XX PN W020012920-A2.

PD 05-APR-2001.

XX 28-SEP-2000: 2000MO-US26524.

XX 29-SEP-1999: 99US-0157137.

PR 03-NOV-1999: 99US-0163280.

XX (HUMA-) HUMAN GENOME SCI INC.

PI Ruben SM, Barash SC, Birse CE, Rosen CA;

DR WPI: 2001-235357/24.

DR P-PSDB: AAG74647.

PT Nucleic acid encoding 4277 human colon cancer-associated polypeptides,  
PS useful for preventing, diagnosing and/or treating colorectal cancers -  
Claim 1; Page 2955; 9803pp: English.

CC AAH32943 to AAH37195 and AAG73514 to AAG77788 represent human colon  
CC cancer-associated nucleic acid molecules (N) and proteins (P), where  
CC the proteins are collectively known as colon cancer antigens. The colon  
CC cancer antigens have cytostatic activity and can be used in gene  
CC therapy and vaccine production. N and P may be used in the prevention,  
CC diagnosis and treatment of diseases associated with inappropriate P  
CC expression. For example, N and P may be used to treat disorders  
CC associated with decreased expression by rectifying mutations or deletions  
CC in a patient's genome that affect the activity of P by expressing P.  
CC Inactive proteins or to supplement the patient's own production of P.  
CC Additionally, N may be used to produce the colon cancer-associated Ps,  
CC by inserting the nucleic acids into a host cell and culturing the cell  
CC to express the proteins. N and P can be used in the prevention, diagnosis  
CC and treatment of colorectal carcinomas and cancers. AAH37196 to AAH37204  
CC and AAB77789 represent sequences used in the exemplification of the  
CC present invention.

CC N.B. Pages 666 to 682 and page 7053 of the sequence listing were  
CC missing at time of publication, meaning no sequences are present for  
CC SEQ ID NO:1027 to 1052, 7921 and 7922.

SO Sequence 806 BP; 163 A; 253 C; 260 G; 128 T; 2 other;

Query Match 62.1%; Score 389; DB 22; Length 806;  
Best Local Similarity 100.0%; Pred. No. 3.2e-61;  
Matches 389; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 43 GGAGCAGAGTGGAGAGAGCGGGAGTCCAGGTTCCGCCCGGAGCGCACTTCTCC 102  
DB 1 GGAGCAGAGTGGAGAGAGCGGGAGTCCAGGTTCCGCCCGGAGCGCACTTCTCC 60  
QY 103 TGGTGGCGGCTGACAGCGGAGTGAAGCGGCGAGCGGGGATCTGAGAGCATGGG 162  
DB 61 TGGTGGCGGCTGACAGCGGAGTGAAGCGGCGAGCGGGGATCTGAGAGCATGGG 120  
QY 163 CGGCGCGGAGCGCATCTCTGATGCGCTGGAAGACCTGACCGCGGAGACTTAAGAG 222  
DB 121 CGCGCGGCGAGCGCATCTCTGATGCGCTGGAAGACCTGACCGCGGAGACTTAAGAG 180  
QY 223 TTCAAGCTGAAGTGTGTGCGGTGCGGTGCGGAGGCTACGGGCGCATCCCGGGG 282  
DB 181 TTCAAGCTGAAGTGTGTGCGGTGCGGTGCGGAGGCTACGGGCGCATCCCGGGG 240  
QY 283 GCGTGTGTCTCATGGAGCGCTTGGACCTACCGACAAAGCTGTAGCTTACTTGGAG 342  
DB 241 GCGTGTGTCTCATGGAGCGCTTGGACCTACCGACAAAGCTGTAGCTTACTTGGAG 300  
QY 343 ACTACGCGCGAGCTACCGCTAAGTGTGCTGCGAGACATGGGCTCAGAGATGGCC 402  
DB 301 ACTACGCGCGAGCTACCGCTAAGTGTGCTGCGAGACATGGGCTCAGAGATGGCC 360  
QY 403 GGGCAGCTGAGGCGGCGACGACAGG 431  
DB 361 GGGCAGCTGAGGCGGCGACGACAGG 389

RESULT 4  
AAC77884  
ID AAC77884 standard; cDNA; 811 BP.  
XX  
XX AAC77884;  
XX  
DT 08-FEB-2001 (first entry)  
DE Human cancer associated gene sequence SEQ ID NO:278.  
XX  
XX Human: cancer associated gene; cancer antigen; detection; cancer;  
diagnosis; cytostatic; proliferative; vulnerrary; immunomodulator;  
antidiabetic; antiaesthetic; antineumatic; antiaesthetic; antiviral;  
antiinflammatory; antihypertoid; antiallergic; antibacterial; cardiac;  
dermatological; neuroprotective; thrombolytic; coagulant; neutropic;  
vasotropic; antiporatic; antigenogenic; gene therapy; inflammation;  
immune disorder; haematopoietic cell disorder; autoimmune disorder;  
allergic reaction; graft versus host disease; organ rejection;  
haemostatic; thrombolytic; cardiovascular disorder; infection;  
neurological disease; drug screening; ss.  
XX  
XX Homo sapiens.  
OS  
XX  
XX MO200055350-A1.  
PN  
XX  
XX 21-SEP-2000.  
BD  
XX  
XX 08-MAR-2000; 2000MO-US05882.  
PF  
XX  
XX 12-MAR-1999; 99US-0124270.  
PR  
XX  
XX (HUMA-) HUMAN GENOME SCI INC.  
XX  
XX Rosen CA, Ruben SM;  
PI  
XX  
XX WPI; 2000-587533/55.  
DR  
XX  
XX P-PSDB; AAB43675.  
DR  
XX  
XX Novel isolated nucleic acids comprising sequences encoding peptides  
PT useful for treating or diagnosing e.g. cancer -  
PT  
XX  
XX Claim 1; Page 841; 2352pp; English.  
PS  
XX  
XX AAC77607 to AAC78448 encode the human cancer associated proteins given  
CC in AAB43398 to AAB44239. The proteins can have activities based on the  
CC tissues and cells the genes are expressed in. Example of activities  
CC include: cytostatic; proliferative; vulnerrary; immunomodulator;  
CC antidiabetic; antiaesthetic; antineumatic; antiaesthetic;  
CC antiinflammatory; antihypertoid; antiallergic; antibacterial;  
CC dermatological; neuroprotective; cardiac; thrombolytic; coagulant;  
CC neutropic; vasotropic; antiporatic and antigenogenic. The  
CC polynucleotides and polypeptides can be used for preventing, treating or  
CC ameliorating medical conditions and diagnosing pathological conditions.  
CC Polynucleotides, polypeptides, antibodies, agonists and antagonists from  
CC the present invention may be used to treat immune disorders by activating  
CC or inhibiting the proliferation, differentiation or mobilisation of  
CC immune cells, to treat disorders of haematopoietic cells, autoimmune  
CC disorders, allergic reactions, graft versus host disease and organ  
CC rejection, modulate haemostatic or thrombolytic activity, modulate  
CC inflammation, cancers, cardiovascular disorders, neurological disease and  
CC bacterial or viral infections. The peptides, nucleotides, antibodies,  
CC agonists and antagonists may be also be used in drug screens. AAC78449 to  
CC AAC78457 and AAB44240 represent sequences used in the exemplification of  
CC the present invention.  
XX  
XX Sequence 811 BP; 170 A; 253 C; 259 G; 128 T; 1 other:  
SO  
Query Match 62.1%; Score 389; DB 21; Length 811;  
Best Local Similarity 100.0%; Pred. No. 3.2e-61;  
Matches 389; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

DB 1 GAGAGACAGAGTGGGAGGAGAGCGGGAGTCCAGGTTCCGCCCGGAGCGACTTCTCC 60  
|||  
QY 103 TGGTGGCGGCTGACAGCGGGGTGAGCGCGGCGGCGGGATCTGAGGCGATGAGG 162  
|||  
DB 61 TGGTGGCGGCTGACAGCGGGGTGAGCGCGGCGGCGGGATCTGAGGCGATGAGG 120  
|||  
QY 163 CGGCGCGCGGAGCGGATCTGAGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 222  
|||  
DB 121 CGGCGCGCGGAGCGGATCTGAGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 180  
|||  
QY 223 TTCAAGCTGAAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 282  
|||  
DB 181 TTCAAGCTGAAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 240  
|||  
QY 283 GGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 342  
|||  
DB 241 GGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 300  
|||  
QY 343 ACCTAGGCGCGCGGAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 402  
|||  
DB 301 ACCTAGGCGCGCGGAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 360  
|||  
QY 403 GGGAGCTGAGCG 431  
|||  
DB 361 GGGAGCTGAGCG 389  
|||  
RESULT 5  
AA23631  
ID AA23631 standard; cDNA; 779 BP.  
XX  
XX  
XX AA23631;  
AC  
XX  
XX 08-DEC-1999 (first entry)  
DT  
XX  
XX Human breast tumour-associated EST 21.  
DE  
XX  
XX Expressed sequence tag; EST; human; breast; cancer; gene therapy;  
KW treatment; tumour; cytostatic; medicament; ss.  
XX  
XX  
XX Homo sapiens.  
OS  
XX  
XX DE19813839-A1.  
PN  
XX  
XX 23-SEP-1999.  
PD  
XX  
XX 20-MAR-1998; 98DE-1013839.  
PF  
XX  
XX 20-MAR-1998; 98DE-1013839.  
PR  
XX  
XX (META-) METAGEN GES GENOMFORSCHUNG MBH.  
PA  
XX  
XX Specht T, Hinzmann B, Schmitt A, Pilarsky C, Dahl E, Rosenthal A;  
PI WPI; 1999-528981/45.  
XX  
XX Human nucleic acid sequences and protein products from tumor breast  
PT tissue, useful for breast cancer therapy -  
PT  
XX  
XX Claim 1a; 101; 188pp; German.  
PS  
XX  
XX This invention describes novel human nucleic acid sequences from tumor  
CC breast tissue which have cytostatic activity. The nucleic acid sequences  
CC can be used to produce and isolate full-length gene sequences. They can  
CC be used to express proteins, which can be used as tools to find an  
CC activity against breast cancer. The sequences can be used in sense or  
CC antisense form. They are especially useful for medicaments for gene  
CC therapy to treat breast cancer. AA23631-248617 represents expressed  
CC sequence tags described in the method of the invention.  
XX  
XX Sequence 779 BP; 149 A; 252 C; 251 G; 127 T; 0 other;

Query Match 58.8%; Score 368; DB 20; Length 779;  
 Best Local Similarity 100.0%; Pred. No. 1.8e-57;  
 Matches 368; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 64 GCGGGAGTCCAGGTTCCGCCCGGAGCCGACTTCTCTGTTGGCGGCTGACGGGG 123  
 DB 1 GCGGGAGTCCAGGTTCCGCCCGGAGCCGACTTCTCTGTTGGCGGCTGACGGGG 60  
 QY 124 TGAGCGGCGGAGCGGCGGCGGATCCTGGAGCCATGGGGCGCGCGCAGCCATCCTG 183  
 DB 61 TGAGCGGCGGAGCGGCGGCGGATCCTGGAGCCATGGGGCGCGCGCAGCCATCCTG 120  
 QY 184 GATGCGCTGGAGAACTGACCCCGAGAGACTCAAGAACTTCAAGCTGAAGCTGCTGCG 243  
 DB 121 GATGCGCTGGAGAACTGACCCCGAGAGACTCAAGAACTTCAAGCTGCTGCG 180  
 QY 244 GTGCGCGTGCAGAGGCTACGGGGCATCCCGGGGGCGCGCTGTCTCATMGAGCGCC 303  
 DB 181 GTGCGCGTGCAGAGGCTACGGGGCATCCCGGGGGCGCGCTGTCTCATMGAGCGCC 240  
 QY 304 TTGAGACTCAGCCGACAGCTGTGCTGACTTCTTACCTGGAGACTACGGCGCCGACTCAC 363  
 DB 241 TTGAGACTCAGCCGACAGCTGTGCTGACTTCTTACCTGGAGACTACGGCGCCGACTCAC 300  
 QY 364 GCTACGCTGCTGCAGCATGGGCTGCAGAGATGGCGGGGAGCTGCAGGGCGGCGACG 423  
 DB 301 GCTACGCTGCTGCAGCATGGGCTGCAGAGATGGCGGGGAGCTGCAGGGCGGCGACG 360  
 QY 424 CACGAGGG 431  
 DB 361 CACGAGGG 368

RESULT 6  
 ID AAL47126 standard; DNA: 782 BP.  
 AC AAL47126;

DT 20-AUG-2002 (first entry)

DE Pyrin domain containing protein Pycard coding sequence.

XX Pyrin domain; PYD domain; antiinflammatory; antiparkinsonian;  
 KW antiarteriosclerotic; antispasmodic; antibacterial; vituocide;  
 KW neuroprotective; antiarthritic; antirheumatic; antisthmatic;  
 KW nephrotropic; osteopachic; nootropic; intracellular signal transduction;  
 KW inflammation; Alzheimer's disease; infection; psoriasis; asthma;  
 KW arteriosclerosis; multiple sclerosis; rheumatoid arthritis; sarcoidosis;  
 KW osteoarthritis; glomerulonephritis; gene; ds.

XX Unidentified.

OS WO200240668-A2.

PN 23-MAY-2002.

PD 30-OCT-2001; 2001WO-EP12545.

PF 15-NOV-2000; 2000DE-1056687.

PR 30-NOV-2000; 2000DE-1059595.

XX (APOT-) APOTEC RES & DEV LTD.

PA Tschoopp J, Martinon F;

XX WPI: 2002-427093/45.

DR P-PSDB; AAO17854.

XX New DNA encoding protein with pyrin domain, useful for treating  
 PT diseases involving impaired signal transduction, particularly  
 PT inflammation, also proteins and antibodies -  
 XX

PS Claim 5; Fig 1; 116pp; German.

XX The present invention relates the DNA and their encoded proteins, where  
 CC the proteins contain at least one PYD (pyrin) domain. These can be used  
 CC to treat diseases associated with impaired intracellular signal  
 CC transduction, particularly inflammation such as psoriasis,  
 CC arteriosclerosis, bacterial or viral infections (particularly meningitis  
 CC and pneumonia), multiple sclerosis, rheumatoid arthritis, asthma,  
 CC sarcoidosis, glomerulonephritis and osteoarthritis, and also Alzheimer's  
 CC and Parkinson's diseases. The present sequence is a coding sequence of  
 CC the invention.

XX Sequence 782 BP; 158 A; 251 C; 246 G; 127 T; 0 other.

Query Match 57.7%; Score 361; DB 24; Length 782;  
 Best Local Similarity 100.0%; Pred. No. 3.3e-56;  
 Matches 361; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 71 GTCCAGGTTCCGCCCGGAGCCGACTTCTCTGTTGGCGGCTGACGGGGTGAACGG 130  
 DB 1 GTCCAGGTTCCGCCCGGAGCCGACTTCTCTGTTGGCGGCTGACGGGGTGAACGG 60  
 QY 131 CGGCGGCGGCGGAGATCCTGGAGCCATGGGGCGCGCGCAGCCATCCTGGATGCCG 190  
 DB 61 CGGCGGCGGCGGAGATCCTGGAGCCATGGGGCGCGCGCAGCCATCCTGGATGCCG 120  
 QY 191 TGAGAACTGACCCCGGAGAGCTCAAGAACTTCAAGCTGAAGCTGCTGTGGTGGCGG 250  
 DB 121 TGAGAACTGACCCCGGAGAGCTCAAGAACTTCAAGCTGAAGCTGCTGTGGTGGCGG 180  
 QY 251 TGGCGGAGGCTACGGGGCATCCCGGGGGCGCGCTGTCTCATMGAGCGCTTGGAGC 310  
 DB 181 TGGCGGAGGCTACGGGGCATCCCGGGGGCGCGCTGTCTCATMGAGCGCTTGGAGC 240  
 QY 311 TCACCGAAGCTGTGCTGACTTCTTACCTGGAGACTACGGCGCCGAGCTCACCGCTAACG 370  
 DB 241 TCACCGAAGCTGTGCTGACTTCTTACCTGGAGACTACGGCGCCGAGCTCACCGCTAACG 300  
 QY 371 TGTGCGCGCATGGGCTGCAGAGATGGCGGCGAGCTGCAGGGCGGCGACGACGAG 430  
 DB 301 TGTGCGCGCATGGGCTGCAGAGATGGCGGCGAGCTGCAGGGCGGCGACGACGAG 360  
 QY 431 G 431  
 DB 361 G 361

RESULT 7  
 ID AAD03906 standard; cDNA: 713 BP.  
 AC AAD03906;

DT 02-JUL-2001 (first entry)

DE Alternatively spliced form of human TMS1 cDNA (lacking exon2).

XX Human; target of methylation-induced silencing-1; TMS1; cytosolic;  
 KW antiproliferative; apoptosis inducer; gene therapy; CpG island;  
 KW caspase-recruiting domain; CARD; cancer; breast; ss.

OS Homo sapiens.

XX Key Location/Qualifiers

FT CDS 75..605

FT /tag= a /product= "Alternatively spliced form of human TMS1  
 FT protein lacking exon2"

PN WO200129235-A2.

XX 26-APR-2001.

[illegible]

DE	Alternative splicing	form of human TMS1 cDNA (lacking exon3).
KW	Human; target of methylation-induced silencing-1; TMS1; cytostatic;	
KW	antiproliferative; apoptosis inducer; gene therapy; CpG island;	
KW	caspase-recruiting domain; CAR; cancer; breast; ss.	
XX		
OS	Homo sapiens.	
XX		
PH	Key	Location/Qualifiers
FT	CDS	75..404
FT		/*tag= a
FT		/product= "alternatively spliced form of human TMS1
FT		protein lacking exon3"
FT		/note= "CDS does not include stop codon"
FT		/partial
PN	MO200129235-A2.	
PD	26-APR-2001.	
XX		
PF	18-OCT-2000; 2000MO-US28747.	
XX		
PR	18-OCT-1999; 99US-0159975.	
XX		
PA	(UYEM-) UNTY EMORY.	
PI	Vertino PM;	
DR	WPI; 2001-290922/30.	
XX	P-PSDB; AAB00595.	
PT	Novel gene TMS1, transcriptionally silenced due to increased	
PT	methylation useful for identifying subject at risk of developing tumor	
PT	characterized by abnormal methylation, for treating cancer by inducing	
PT	apoptosis -	
PS	Claim 106; Page 123-124; 124pp; English.	
XX		
XX	The invention relates to identification of target of methylation-induced	
CC	silencing-1 (TMS1) gene. This gene is transcriptionally silenced due to	
CC	abnormal methylation of a CpG island in its 5' regulatory region. TMS1	
CC	consists of a carboxy terminal caspase-recruiting domain (CAR) and	
CC	plays a role in induction of apoptosis. TMS1 gene and protein are useful	
CC	as tools for diagnosing and treating a subject at risk of developing	
CC	cancer (e.g. breast cancer) characterized by abnormal CpG methylation or	
CC	abnormally low levels of TMS1 expression products. Unique fragments of	
CC	TMS1 gene are used as probes. TMS1 gene is useful in gene therapy.	
CC	TMS1 molecule is also useful for treating abnormal cell proliferation by	
CC	increasing TMS1 polypeptide level to an above normal level. The CpG	
CC	island region of TMS1 or its fragments are used to study the methylation	
CC	patterns apart from any coding region contained in it.	
CC	The present sequence is alternatively spliced form of human target of	
CC	methylation-induced silencing-1 (TMS1) cDNA lacking exon3.	
CC	Note: The present sequence designated as SEQ ID NO:26, is referred	
CC	as DNA throughout the specification, however this sequence has been	
CC	referred as protein in claim 106.	
XX		
SO	Sequence 405 BP; 63 A; 139 C; 147 G; 56 T; 0 other;	
Query Match	54.78; Score 342.6; DB 22; Length 405;	
Best Local Similarity	98.9%; Pred. Mis. 6.0e-53;	
Matches 345; Conservative	0; Mismatches 4; Indels 0; Gaps 0;	
OY	83 CCCCAGAGCGCACTCTCTCTGTCGGCGCGTCAGCGGGGTAGCGCGAGCGCGC 142	
DB	1 CCACGGGTCCTCCCTCTCTGTCGGCGCGTCAGCGGGGTAGCGCGAGCGCGC 60	
OY	143 GGGATCTTGGAGCCATGGGGCGCGCGCGAGCCCATCTCGATGCGCTGGAGAACTGA 202	
DB	61 GGGATCTTGGAGCCATGGGGCGCGCGCGAGCCCATCTCGATGCGCTGGAGAACTGA 120	
OY	203 CCCCAGAGCGCACTCTCTCTGTCGGCGCGTCAGCGGGGTAGCGCGAGCGCGCT 262	



Db 121 CCGCGGAGAGCTCAAGAAGTTCAAGCTGAGCTGCTGCTGCGCGCTGCGGAGGCT 180  
Qy 263 ACGGGGCGCATCCCCGGGGGGCGCGTGTCTCATGAGACCTTGGACTCAGCAGAC 322  
Db 181 ACGGGGCGCATCCCCGGGGGGCGCGTGTCTCATGAGACCTTGGACTCAGCAGAC 240  
Qy 323 TGGTCAGCTTCTACCTGGAGACCTACGCGCGCGAGCTACCGCTTAACGCTGCGGAC 382  
Db 241 TGGTCAGCTTCTACCTGGAGACCTACGCGCGCGAGCTCAGCAGCTTAACGCTGCGGAC 300  
Qy 383 TGGGCGCTGAGAGAGATGGCGCGGCACTGCAAGCGCGCCACGACCAAGG 431  
Db 301 TGGGCGCTGAGAGAGATGGCGCGGCACTGCAAGCGCGCCACGACCAAGG 349

RESULT 9  
AAD03890  
ID AAD03890 standard; cDNA: 770 BP.

AC AAD03890;

DT 02-JUL-2001 (first entry)

DE Human target of methylation-induced silencing-1 (TMS1) cDNA.

KW Human: target of methylation-induced silencing-1; TMS1; cytosstatic;  
antiproliferative; apoptosis inducer; gene therapy; CpG island;  
caspase-recruiting domain; CARD; cancer; breast; ss.

OS Homo sapiens.

XX Key Location/Qualifiers

FT CDS 75..662

FT /tag= a

FT /product= "Human TMS1 protein"

PM WO200129235-A2.

XX 26-APR-2001.

PF 18-OCT-2000; 2000WO-US28747.

PR 18-OCT-1999; 99US-0159975.

PA (UYEM-) UNIV EMORY.

PI Vertino PM;

DR WPI: 2001-290922/30.

DR P-PSDB; AAED0588.

PT Novel gene TMS1, transcriptionally silenced due to increased  
methylation useful for identifying subject at risk of developing tumor  
characterized by abnormal methylation, for treating cancer by inducing  
apoptosis -

PS Claim 68; Page 113-114; 124pp; English.

XX The invention relates to identification of target of methylation-induced  
silencing-1 (TMS1) gene. This gene is transcriptionally silenced due to  
abnormal methylation of a CpG island in its 5' regulatory region. TMS1  
consists of a carboxy terminal caspase-recruiting domain (CARD) and  
plays a role in induction of apoptosis. TMS1 gene and protein are useful  
as tools for diagnosing and treating a subject at risk of developing  
cancer (e.g. breast cancer) characterized by abnormal CpG methylation or  
abnormally low levels of TMS1 expression products. Unique fragments of  
TMS1 gene are used as probes. TMS1 gene is useful in gene therapy.  
XX TMS1 molecule is also useful for treating abnormal cell proliferation by  
increasing TMS1 polypeptide level to an above normal level. The CpG  
island region of TMS1 or its fragments are used to study the methylation  
patterns apart from any coding region contained in it.  
CC The present sequence is human target of methylation-induced silencing-1  
(TMS1) cDNA.

XX SQ Sequence 770 BP; 157 A; 247 C; 241 G; 125 T; 0 other;

Query Match 54.7%; Score 342.6; DB 22; Length 770;

Best Local Similarity 98.9%; Pred. No. 6.5e-33;

Matches 345; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

Qy 83 CCGCGAGAGCTTCCCTGCTGCGGCGGCTGACGCGGAGTGAGCGGCGGCGGCGG 142

Db 1 CCACGCTGCGACTTCTCTGCTGCGGCGGCTGACGCGGAGTGAGCGGCGGCGGCGG 60

Qy 143 GGGATCTTGGAGACCTAGGCGCGCGCGGAGGCGCTTGGATGCGCTGAGAACCTGA 202

Db 61 GGGATCTTGGAGACCTAGGCGCGCGCGGAGGCGCTTGGATGCGCTGAGAACCTGA 120

Qy 203 CCGCGAGAGCTCAAGAAGTTCAAGCTGAGCTGCTGCTGCGCGCTGCGGAGGCT 262

Db 121 CCGCGAGAGCTCAAGAAGTTCAAGCTGAGCTGCTGCTGCGCGCTGCGGAGGCT 180

Qy 263 ACGGGGCGCATCCCCGGGGGGCGCGTGTCTCATGAGACCTTGGACTCAGCAGAC 322

Db 181 ACGGGGCGCATCCCCGGGGGGCGCGTGTCTCATGAGACCTTGGACTCAGCAGAC 240

Qy 323 TGGTCAGCTTCTACCTGGAGACCTACGCGCGCGAGCTACCGCTTAACGCTGCGGAC 382

Db 241 TGGTCAGCTTCTACCTGGAGACCTACGCGCGCGAGCTACCGCTTAACGCTGCGGAC 300

Qy 383 TGGGCGCTGAGAGAGATGGCGCGGCACTGCAAGCGCGCCACGACCAAGG 431

Db 301 TGGGCGCTGAGAGAGATGGCGCGGCACTGCAAGCGCGCCACGACCAAGG 349

RESULT 10  
AAD03892  
ID AAD03892 standard; DNA: 340 BP.

AC AAD03892;

DT 02-JUL-2001 (first entry)

DE Human target of methylation-induced silencing-1 (TMS1) exon1 DNA.

KW Human: target of methylation-induced silencing-1; TMS1; cytosstatic;  
antiproliferative; apoptosis inducer; gene therapy; CpG island;  
caspase-recruiting domain; CARD; cancer; breast; ds.

XX Homo sapiens.

XX Key Location/Qualifiers

FT CDS 67..339

FT /tag= a

FT /product= "Human TMS1 exon1 protein"

PM WO200129235-A2.

XX 26-APR-2001.

PF 18-OCT-2000; 2000WO-US28747.

PR 18-OCT-1999; 99US-0159975.

PA (UYEM-) UNIV EMORY.

PI Vertino PM;

DR WPI: 2001-290922/30.

DR P-PSDB; AAED0589.

PT Novel gene TMS1, transcriptionally silenced due to increased  
methylation useful for identifying subject at risk of developing tumor  
characterized by abnormal methylation, for treating cancer by inducing



Db 307 TGCAGGCGCACGCCACGAGG 328

RESULT 12

ABK87966

ID ABK87966 standard; cDNA; 740 BP.

XX

XX ABK87966;

XX

DT 07-OCT-2002 (first entry)

XX

DE Human caspase recruitment domain-5 (CARD-5) cDNA.

XX

KW Human; gene; ss; caspase recruitment domain-5; CARD-5; antiinflammatory;

KW immunosuppressive; caspase; cysteinyl aspartate-specific proteinase;

KW apoptosis; nuclear factor-kappa B; NF-kappaB; transcription factor;

KW cell proliferation; gene therapy; immune disorder;

KW chronic inflammatory disease; Hashimoto's thyroiditis; graft rejection;

KW sarcoidosis; atopy; asthma; allergy; glomerular nephritis;

KW human immunodeficiency virus; HIV; bacterial infection; tuberculosis;

KW lepromatous leprosy; autoimmune disorder; systemic lupus erythematosus;

KW arthritis; cell depletion; neurological disorder; Alzheimer's disease;

KW Parkinson's disease; spinal muscular atrophy; haematologic disease;

KW myelodysplastic syndrome; aplastic anaemia; myocardial infarction;

KW stroke.

XX

OS Homo sapiens.

XX

XX Key Location/Qualifiers

XX CDS 54..641

XX FT /\*tag= a

XX FT /product= "CARD-5"

XX

XX PN W0200244354-A2.

XX

XX PD 06-JUN-2002.

XX

XX PE 29-NOV-2001; 2001WO-US44894.

XX

XX PR 01-DEC-2000; 2000US-0728721.

XX PR 24-APR-2001; 2001US-0841879.

XX

XX PA (MILL-) MILLENIUM PHARM INC.

XX

XX PI Bertin J;

XX

XX DR WPI: 2002-557538/59.

XX DR P-PSDB; AAD99353.

XX

XX PT Novel isolated murine or human caspase recruitment domain (CARD)-5

XX PT polypeptide, useful for treating immune disorders such as Hashimoto's

XX PT thyroiditis, graft rejection, allergy, glomerular nephritis,

XX PT tuberculosis

XX

XX PS Claim 7; Fig 3; 100pp; English.

XX

XX The invention discloses the isolated polypeptides, and encoding nucleic

XX acids, of murine and human caspase recruitment domain (CARD)-5. Caspases

XX (cysteinyl aspartate-specific proteinases) are central to the apoptotic

XX program and responsible for the degradation of cellular proteins that

XX lead to the morphological changes seen in cells undergoing apoptosis.

XX CC Caspases interact with other caspases via their CARs and different

XX CC subtypes of CARs may confer binding specificity. CARD-5 is an

XX CC intracellular protein that is predicted to be involved in regulating

XX CC caspase activation. CARD-5 activates the nuclear factor-kappa B

XX CC (NF-kappaB) transcription factor pathway and binds the CARs of

XX CC caspase-1, CARD-7 and itself. CARD-5 can, therefore, modulate CARD-5

XX CC activity and NF-kappaB activation, regulate cell growth and cell death

XX CC and be used in gene therapy. The CARD-5 polypeptides are useful for

XX CC identifying compounds which bind to them and modulates their activity and

XX CC for detecting the presence of CARD-5 in a sample. CARD-5 polypeptides,

XX CC nucleic acids, antibodies and modulators of CARD-5 expression or activity

XX CC can be used to treat immune disorders such as chronic inflammatory

CC diseases and disorders, Hashimoto's thyroiditis, graft rejection,

CC sarcoidosis, atopic conditions (such as asthma and allergy), glomerular

CC nephritis, human immunodeficiency virus (HIV) and bacterial infections

CC (including tuberculosis and lepromatous leprosy) and in screening and

CC detection assays. Modulators of CARD-5 activity or expression are also

CC useful for treating autoimmune disorders, such as systemic lupus

CC erythematosus and arthritis, cell depletion, neurological disorders,

CC such as Alzheimer's disease, Parkinson's disease and spinal muscular

CC atrophy, haematologic diseases, such as myelodysplastic syndrome and

CC aplastic anaemia, myocardial infarction and stroke. The sequence

CC presented is the human caspase recruitment domain-5 (CARD-5) cDNA.

XX

SQ Sequence 740 BP; 146 A; 238 C; 236 G; 120 T; 0 other;

XX

Query Match 51.4%; Score 322; DB 24; Length 740;

Best Local Similarity 100.0%; Pred. No. 3.2e-49;

Matches 322; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 110 GCGCTGCAGCGGGGTAGCGGCGCGCGCGCGGCGATCCTGGAGCCATGGGCGCGGC 169

DB 7 CGCGTCGACGCGGGGTAGCGGCGCGCGCGCGCGGCGATCCTGGAGCCATGGGCGCGGC 66

QY 170 GCGAGCCATCTCTGGATGCGCTGGAGAACTGACCGCGCGAGCTCAAGAGTTCAAGC 229

DB 67 GCGAGCCATCTCTGGATGCGCTGGAGAACTGACCGCGCGAGCTCAAGAGTTCAAGC 126

QY 230 TGAAGCTGCTGCGGTGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 289

DB 127 TGAAGCTGCTGCGGTGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 186

QY 290 TGTCCATGAGAGCGCTTGGACCTGACCGAAGCTGTGACCTTCTACCTGGAGACTACG 349

DB 187 TGTCCATGAGAGCGCTTGGACCTGACCGAAGCTGTGACCTTCTACCTGGAGACTACG 246

QY 350 GCGCGAGCTCACCGCTAACGTGCTGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 409

DB 247 GCGCGAGCTCACCGCTAACGTGCTGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 306

QY 410 TGCAGGCGGCCACGACGAGG 431

DB 307 TGCAGGCGGCCACGACGAGG 328

RESULT 13

ABK87967/C

ID ABK87967 standard; DNA; 740 BP.

XX

XX AC ABK87967;

XX

XX DT 07-OCT-2002 (first entry)

XX

XX DE Human caspase recruitment domain-5 (CARD-5) cDNA complementary strand.

XX

XX Human; ss; caspase recruitment domain-5; CARD-5; antiinflammatory;

XX immunosuppressive; caspase; cysteinyl aspartate-specific proteinase;

XX apoptosis; nuclear factor-kappa B; NF-kappaB; transcription factor;

XX cell proliferation; gene therapy; immune disorder;

XX chronic inflammatory disease; Hashimoto's thyroiditis; graft rejection;

XX sarcoidosis; atopy; asthma; allergy; glomerular nephritis;

XX human immunodeficiency virus; HIV; bacterial infection; tuberculosis;

XX lepromatous leprosy; autoimmune disorder; systemic lupus erythematosus;

XX arthritis; cell depletion; neurological disorder; Alzheimer's disease;

XX Parkinson's disease; spinal muscular atrophy; haematologic disease;

XX myelodysplastic syndrome; aplastic anaemia; myocardial infarction;

XX stroke.

XX

XX OS Homo sapiens.

XX

XX PN W0200244354-A2.

XX

XX PD 06-JUN-2002.

XX

XX PF 29-NOV-2001; 2001WO-US44894.



CC thrombocytopenia, stem cell disorders, aplastic anaemia, for  
CC regeneration of bone, cartilage, tendon, ligament and/or nerve tissue  
CC growth, and in tissue repair, healing of burns, incisions, ulcers, for  
CC treating osteoporosis, osteoarthritis, bone degenerative disorders, or  
CC periodontal disease, lung or liver fibrosis, reperfusion injury in  
CC various tissues, various immune deficiencies and disorders including  
CC severe combined immunodeficiency (SCID), bacterial or fungal infections,  
CC autoimmune disorders (e.g. multiple sclerosis, rheumatoid arthritis,  
CC diabetes mellitus, myasthenia gravis), allergic reactions and conditions,  
CC such as asthma or other respiratory problems, coagulation disorders,  
CC haemophilia), septic shock, sepsis, arthritis, nephritis and inflammatory  
CC bowel disease, viral infection and are useful in altering bodily  
CC characteristics. The present sequence encodes a novel protein of the  
CC invention.

XX Sequence 745 BP; 157 A; 234 C; 235 G; 119 T; 0 other;

Query Match 51.3%; Score 321; DB 22; Length 745;

Best Local Similarity 100.0%; Pred. No. 4.8e-49;

Matches 321; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 111 GCGTCAGCGGGGTGACGGCGGCGGATCCTGAGCCATGGGGCGCGCG 170

Db 2 GCGTCAGCGGGGTGACGGCGGCGGATCCTGAGCCATGGGGCGCGCG 61

QY 171 CGACGCCATCCTGATGCGCTGAGAACCTGACCCCGAGAGCTCAAGATTCAAGCT 230

Db 62 CGACGCCATCCTGATGCGCTGAGAACCTGACCCCGAGAGCTCAAGATTCAAGCT 121

QY 231 GAAGCTGCTGTGCTGCGCGCTGCGGAGGCTACGGGGCATCCCGGGGCGCGCTGCT 290

Db 122 GAAGCTGCTGTGCTGCGCGCTGCGGAGGCTACGGGGCATCCCGGGGCGCGCTGCT 181

QY 291 GTCATGAGCGGCTTGACCTCAACCAAGCTGTGAGCTTCTACTGAGAGACTTACG 350

Db 182 GTCATGAGCGGCTTGACCTCAACCAAGCTGTGAGCTTCTACTGAGAGACTTACG 241

QY 351 CGCCGAGCTCAACCGCTTAACGTGTGCGCGGAGCATGGGCTTGCAGAGAGATGGCCGGGACGCT 410

Db 242 CGCCGAGCTCAACCGCTTAACGTGTGCGCGGAGCATGGGCTTGCAGAGAGATGGCCGGGACGCT 301

QY 411 GCAGGGGCGCACGACACGAGG 431

Db 302 GCAGGGGCGCACGACACGAGG 322

RESULT 15

AAC98638 AAC98638 standard; cDNA; 432 BP.

AC AAC98638;

XX 09-MAR-2001 (first entry)

DE Human colon cancer antigen nucleotide sequence SFO ID NO:648.

XX Human; colon cancer; colon cancer antigen; diagnosis; detection;

KW identification; cytostatic; cardioactive; neuroprotective; vulnerary;

KW immunomodulatory; muscular; gynaecological; gastrointestinal;

KW nephrotropic; antiinfective; antibacterial; gene therapy; wound;

KW neural disorder; immune system disorder; muscular disorder;

KW reproductive disorder; gastrointestinal disorder; renal disorder;

KW infectious disease; cardiovascular disorder; ss.

XX Homo sapiens.

OS MO200055351-A1.

XX 21-SEP-2000.

XX 08-MAR-2000; 2000WO-US05883.

XX 12-MAR-1999; 99US-0124270.

XX (HUMA-) HUMAN GENOME SCI INC.

PA Rosen CA, Ruben SM;

XX WPI: 2000-587534/55.

DR P-PSDB: AAB53881.

XX Colon cancer associated gene sequences, referred to as colon cancer

PT antigens, useful for the treatment, prevention, and diagnosis of colon

PS disorders such as colon cancer -

PS Claim 1; Page 1175; 2104bp; English.

CC AAC97991 to AAC98763 encode the human colon cancer associated proteins,  
CC called human colon cancer antigens, given in AAB5334 to AAB54006. The  
CC human colon cancer antigens can have cytostatic, cardioactive, muscular,  
CC neuroprotective, immunomodulatory, gynaecological, gastrointestinal,  
CC vulnerary, nephrotropic, antiinfective and antibacterial activities, and  
CC can be used in gene therapy. The colon cancer antigen polynucleotides,  
CC proteins and antibodies to the proteins are useful for the prevention,  
CC treatment and diagnosis of colon disorders, such as colon cancer. The  
CC polynucleotides may be used in diagnostics and research, such as for  
CC chromosome identification, and as hybridisation probes. The proteins  
CC may also be used to prevent diseases such as neural disorders, immune  
CC system disorders, muscular disorders, reproductive disorders, infectious  
CC gastrointestinal disorders, wounds, renal disorders, infectious  
CC diseases, and cardiovascular disorders. AAC98764 to AAC98772 and  
CC AAB54007 represent sequences used in the exemplification of the present  
CC invention.

XX Sequence 432 BP; 77 A; 142 C; 152 G; 61 T; 0 other;

Query Match 51.1%; Score 320; DB 21; Length 432;

Best Local Similarity 100.0%; Pred. No. 7.4e-49;

Matches 320; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 112 GCTGACGCGGGGTGACGGCGGCGGATCCTGAGCCATGGGGCGCGCGC 171

Db 8 GCTGACGCGGGGTGACGGCGGCGGATCCTGAGCCATGGGGCGCGCGC 67

QY 172 GACGCCATCCTGATGCGCTGAGAACCTGACCCCGAGAGCTCAAGATTCAAGCTG 231

Db 68 GACGCCATCCTGATGCGCTGAGAACCTGACCCCGAGAGCTCAAGATTCAAGCTG 127

QY 232 AAGCTGCTGTGCTGCGCGCTGCGGAGGCTACGGGCGCATCCCGGGGCGCGCTGCTG 291

Db 128 AAGCTGCTGTGCTGCGCGCTGCGGAGGCTACGGGCGCATCCCGGGGCGCGCTGCTG 187

QY 292 TCCATGAGCGGCTTGACCTCAACCAAGCTGTGAGCTTCTACTGAGAGACTTACG 351

Db 188 TCCATGAGCGGCTTGACCTCAACCAAGCTGTGAGCTTCTACTGAGAGACTTACG 247

QY 352 GCCGAGCTCAACCGCTTAACGTGTGCGCGGAGCATGGGCTTGCAGAGAGATGGCCGGGACGCTG 411

Db 248 GCCGAGCTCAACCGCTTAACGTGTGCGCGGAGCATGGGCTTGCAGAGAGATGGCCGGGACGCTG 307

QY 412 CAGGCGGCGCACGACACGAGG 431

Db 308 CAGGCGGCGCACGACACGAGG 327

Search completed: December 22, 2002, 05:27:28

Job time : 263 secs



GenCore version 5.1.3  
Copyright (c) 1993 - 2002 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: December 22, 2002, 05:24:11 ; Search time 52 Seconds  
(without alignments)  
3691.916 Million cell updates/sec

Title: US-09-691-763B-4

Sequence: 1 agcccttgcgcgagggc.....ggaagggaaggatcatt 626

Scoring table: IDENTITY\_NUC

Gapop 10.0 , Gapext 1.0

Searched: 441362 seqs, 1533381 residues

Total number of hits satisfying chosen parameters: 882724

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%

Listing first 45 summaries

Database :  
1: /cgn2\_6/ptodata1/lna/5A.COMB.seq:\*  
2: /cgn2\_6/ptodata1/lna/5B.COMB.seq:\*  
3: /cgn2\_6/ptodata1/lna/5A.COMB.seq:\*  
4: /cgn2\_6/ptodata1/lna/5B.COMB.seq:\*  
5: /cgn2\_6/ptodata1/lna/PCITUS.COMB.seq:\*  
6: /cgn2\_6/ptodata1/lna/backfile1.seq:\*

Pred. No. is the number of results predicted by chance to have a  
score greater than or equal to the score of the result being printed,  
and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	52.6	8.4	44377	2 US-08-804-227C-7	Sequence 7, Appl
2	52.6	8.4	44377	2 US-08-804-198-1	Sequence 1, Appl
3	51.6	8.2	33529	4 US-09-144-085-3	Sequence 3, Appl
4	51.2	8.2	2943	1 US-08-042-747A-7	Sequence 7, Appl
C 5	50.8	8.1	30001	1 US-08-125-468-1	Sequence 1, Appl
C 6	50.8	8.1	30001	2 US-08-474-933-1	Sequence 1, Appl
7	50	8.0	2244	1 US-08-476-519-10	Sequence 10, Appl
8	50	8.0	2244	5 PCT-US95-09323-10	Sequence 10, Appl
9	50	8.0	2334	1 US-08-476-519-1	Sequence 1, Appl
10	50	8.0	2334	5 PCT-US95-09323-1	Sequence 1, Appl
11	49.4	7.9	1765	1 US-08-459-586-15	Sequence 15, Appl
12	49.4	7.9	1765	2 US-08-282-696-15	Sequence 15, Appl
13	49.4	7.9	18994	1 US-08-459-586-4	Sequence 4, Appl
14	49.4	7.9	18994	2 US-08-282-696-4	Sequence 4, Appl
15	48	7.7	530	3 US-08-758-662-4	Sequence 4, Appl
16	48	7.7	2310	3 US-09-036-987A-25	Sequence 25, Appl
17	48	7.7	2310	4 US-09-370-700-25	Sequence 25, Appl
18	48	7.7	4524	2 US-08-845-998-7	Sequence 7, Appl
19	48	7.7	4524	3 US-09-206-537-7	Sequence 7, Appl
20	48	7.7	4524	2 US-09-430-854-7	Sequence 7, Appl
21	47.4	7.6	800	5 PCT-US95-04801-4	Sequence 4, Appl
C 22	47.4	7.6	50937	4 US-09-428-517-1	Sequence 14, Appl
C 23	46.6	7.4	1656	4 US-09-385-028-14	Sequence 1, Appl
C 24	46.6	7.4	1809	1 US-08-455-001-1	Sequence 1, Appl
25	46.6	7.4	1809	5 PCT-US95-11869-1	Sequence 1, Appl
C 26	46.6	7.4	15079	4 US-09-385-028-1	Sequence 1, Appl
27	46.4	7.4	4188	4 US-07-751-891B-2	Sequence 2, Appl

28	46.4	7.4	4242	3 US-07-705-490-2	Sequence 2, Appl
29	46.2	7.4	861	2 US-08-997-080-97	Sequence 97, Appl
30	46.2	7.4	861	2 US-08-997-362-97	Sequence 97, Appl
31	46.2	7.4	861	3 US-08-873-970-97	Sequence 97, Appl
32	46.2	7.4	861	4 US-09-095-855-97	Sequence 97, Appl
33	46.2	7.4	861	4 US-09-324-542-97	Sequence 97, Appl
34	46.2	7.4	861	4 US-09-205-426-97	Sequence 97, Appl
35	46	7.3	1506	4 US-09-158-767-7	Sequence 7, Appl
36	46	7.3	2261	4 US-09-158-767-1	Sequence 1, Appl
37	45.8	7.3	11219	1 US-07-642-734C-1	Sequence 1, Appl
38	45.8	7.3	11219	3 US-08-439-009A-1	Sequence 1, Appl
39	45.6	7.3	1227	4 US-09-385-028-23	Sequence 23, Appl
40	45.6	7.3	15079	4 US-08-916-917-1	Sequence 1, Appl
C 41	45.2	7.2	4692	2 US-08-916-917-1	Sequence 1, Appl
C 42	45.2	7.2	4692	2 US-08-972-631-1	Sequence 1, Appl
C 43	45.2	7.2	4692	2 US-08-972-629-1	Sequence 1, Appl
C 44	45.2	7.2	4692	2 US-08-972-630-1	Sequence 1, Appl
C 45	45.2	7.2	4692	2 US-08-672-211-1	Sequence 1, Appl

#### ALIGNMENTS

RESULT 1  
US-08-804-227C-7  
; Sequence 7, Application US/08804227C  
; Patent No. 5876991  
; GENERAL INFORMATION:  
; APPLICANT: Dehoff, Bradley S.  
; APPLICANT: Kuntzoss, Stuart A.  
; APPLICANT: Rostock, Paul R., Jr.  
; APPLICANT: Sutton, Kimberly L.  
; TITLE OF INVENTION: POLYPEPTIDE SYNTHASE GENES  
; NUMBER OF SEQUENCES: 15  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: THOMAS G. PLANT 1501  
; STREET: LILLY CORPORATE CENTER  
; CITY: INDIANAPOLIS  
; STATE: IN  
; COUNTRY: USA  
; ZIP: 46285  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM compatible  
; OPERATING SYSTEM: MS-DOS  
; SOFTWARE: ASCII(DOS) text only  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/804, 227C  
; FILING DATE: February 21, 1997  
; CLASSIFICATION: 435  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Plant, Thomas, G.  
; REGISTRATION NUMBER: 35,784  
; TELEPHONE: 317-276-2459  
; TELECOMMUNICATION INFORMATION:  
; INFORMATION FOR SEQ ID NO: 7:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 44377 base pairs  
; TYPE: nucleic acid  
; STRANDEDNESS: single  
; TOPOLOGY: linear  
; MOLECULE TYPE: DNA (genomic)  
; FEATURE:  
; NAME/KEY: CDS  
; LOCATION: 350..14002  
; FEATURE:  
; NAME/KEY: CDS  
; LOCATION: 14046..20036  
; FEATURE:  
; NAME/KEY: CDS  
; LOCATION: 20110..31284  
; FEATURE:

NAME/KEY: CDS  
LOCATION: 31329..36071  
FEATURE:  
NAME/KEY: CDS  
LOCATION: 36155..41830  
US-08-804-227C-7

Query Match 8.4%; Score 52.6; DB 2; Length 44377;  
Best Local Similarity 55.9%; Pred. No. 0.038;  
Matches 100; Conservative 0; Mismatches 79; Indels 0; Gaps 0;

Qy 329 GCTTACCTGAGAGCTACGGCGGAGCTACCGCTAACGTGTCGCGGACATGAGCC 388  
Db 16843 GCATGACCTCCCGACCTACCCCTTCCAAACAGACCGCTACGTGGCGCCGCCGCTCC 16902  
Qy 389 TGCAGAGATGGCCGCGAGCTGACGAGCGCGGACGACGAGGTGAGCCGCCGCTCC 448  
Db 16903 GCGCCAGAGGAGGAGGCTGTGCGGCGGCGGTCTGCGGAGGTGAGACACCCCTGCTCA 16962  
Qy 449 CCTCCACCCGCTTTCCTCCCTCCACCCACACGAGCGCTTACCCCGGCGCTTCGCT 507  
Db 16963 CCGCGCGCTGGAACCTGCCCGGACCGACGAGGTGTGAGCCGCGCATATCCGCT 17021

## RESULT 2

US-08-804-198-1  
Sequence 1, Application US/08804198  
Patent No. 5945320

GENERAL INFORMATION:  
APPLICANT: Burgett, Stanley G.  
APPLICANT: Kubatoss, Stuart A.  
APPLICANT: Rao, Nagaraja R.  
APPLICANT: Richardson, Mark A.  
APPLICANT: Rostock, Paul R., Jr.  
TITLE OF INVENTION: PLANTENOLIDE SYNTHASE GENE  
NUMBER OF SEQUENCES: 6  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: PAUL R. CANTRELL, 1138  
STREET: LILLY CORPORATE CENTER  
CITY: INDIANAPOLIS  
STATE: IN  
COUNTRY: USA  
ZIP: 46285  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: Macintosh  
OPERATING SYSTEM: Macintosh 7.0  
SOFTWARE: Microsoft Word 5.1  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/804,198  
FILING DATE:  
CLASSIFICATION: 435  
ATTORNEY/AGENT INFORMATION:  
NAME: CANTRELL, PAUL R.  
REGISTRATION NUMBER: 36,470  
REFERENCE/DOCKET NUMBER: P9113  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 317-276-3885  
INFORMATION FOR SEQ ID NO: 1:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 44377 base pairs  
TYPE: nucleic acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
MOLECULE TYPE: DNA (genomic)  
FEATURE:  
NAME/KEY: CDS  
LOCATION: 350..14002  
NAME/KEY: CDS  
LOCATION: 14046..20036  
FEATURE:  
NAME/KEY: CDS

LOCATION: 20110..31284  
FEATURE:  
NAME/KEY: CDS  
LOCATION: 31329..36071  
FEATURE:  
NAME/KEY: CDS  
LOCATION: 36155..41830  
US-08-804-198-1

Query Match 8.4%; Score 52.6; DB 2; Length 44377;  
Best Local Similarity 55.9%; Pred. No. 0.038;  
Matches 100; Conservative 0; Mismatches 79; Indels 0; Gaps 0;

Qy 329 GCTTACCTGAGAGCTACGGCGGAGCTACCGCTAACGTGTCGCGGACATGAGCC 388  
Db 16843 GCATGACCTCCCGACCTACCCCTTCCAAACAGACCGCTACGTGGCGCCGCCGCTCC 16902  
Qy 389 TGCAGAGATGGCCGCGAGCTGACGAGCGCGGACGACGAGGTGAGCCGCCGCTCC 448  
Db 16903 GCGCCAGAGGAGGAGGCTGTGCGGCGGCGGTCTGCGGAGGTGAGACACCCCTGCTCA 16962  
Qy 449 CCTCCACCCGCTTTCCTCCCTCCACCCACACGAGCGCTTACCCCGGCGCTTCGCT 507  
Db 16963 CCGCGCGCTGGAACCTGCCCGGACCGACGAGGTGTGAGCCGCGCATATCCGCT 17021

## RESULT 3

US-09-144-085-3  
Sequence 3, Application US/09144085  
Patent No. 6280999

GENERAL INFORMATION:  
APPLICANT: Gustafsson, Claes  
APPLICANT: Bellach, Mary C.  
APPLICANT: Ashley, Gary  
APPLICANT: Sullivan, Bryan  
APPLICANT: Ziemann, Rainer  
TITLE OF INVENTION: SORANGIUM POLYKETIDE SYNTHASES AND ENCODING DNA  
FILE REFERENCE: 30062-20020.20  
CURRENT APPLICATION NUMBER: US/09/144,085  
CURRENT FILING DATE: 1998-08-31  
EARLIER APPLICATION NUMBER: 09/010,809  
EARLIER FILING DATE: 1998-01-22  
NUMBER OF SEQ ID NOS: 8  
SOFTWARE: Patentln Ver. 2.0  
SEQ ID NO 3  
TYPE: DNA  
LENGTH: 33529  
ORGANISM: Sorangium cellulosum  
US-09-144-085-3

Query Match 8.2%; Score 51.6; DB 4; Length 33529;  
Best Local Similarity 47.9%; Pred. No. 0.059;  
Matches 179; Conservative 0; Mismatches 194; Indels 1; Gaps 1;

Qy 41 CGGAGACAGAGTGGAGGAAGCGGAGTCCAGTTCCCGCGGAGCCGACTTCCT 100  
Db 13930 CTGTGACCGCGGAGGAGGAGTGTGACGCGTGGCGGCGACACTGTGTGCGGCG 13989  
Qy 101 CTTGTGCGGCGCTCCAGCGGAGTGAAGCGCGGCGGCGCGG-6GATCTGAGCCATG 159  
Db 13990 CACGGGAGTGGACCTTGTCTGACGTGCGGCGGCGGTGAGGCGCCCGGCGCGC 14049  
Qy 160 GGGCGGCGCGGAGCGCATCTGATGCTCTGAGACCTGACCGCGGAGAGCTCAAG 219  
Db 14050 GAGCTGTGTGATTCGCTGCGGAGCTGCGGCGCGGAGACCGTACGCTGCGGTGAGC 14109  
Qy 220 AAGTTACAGCTAAGCTCTGTCGTCGCGCGGAGGAGGCTTACGGCGCATCCGCGG 279  
Db 14110 GTGTGAGAGGAGAGGTGCGCGTGTGCTGCGCGGAGTATGAGCGCGCGCGCTG 14169  
Qy 280 GCGCGCTGCTTCATGAGAGCGCTTGAACCTCACCGACAGCTGTGAGCTTCTACCTG 339



Db 14170 AGCGGCTGCTGCACTGGCCGCGCTCGACGAGCGCTGTCGCCGCCAGACGCC 14229  
QY 340 GAGACCTAGCGGGCGGCGGCTACCGGTAAGCTGAGCGGAGCGGCGGCGAGGATG 399  
Db 14230 GAGCGCTCTGCGGGGCTGCTGGCCGGAAGTGGAGCGGGGCGCTGCACTGCGAGCTG 14289  
QY 400 GCGGCGCACTGCA 413  
Db 14290 AC CGCGGAGCTGGA 14303

RESULT 4  
US-08-042-747A-7  
; Sequence 7, Application US/08042747A  
; Patent No. 5487969  
; GENERAL INFORMATION:  
; APPLICANT: Eberle, Richard  
; APPLICANT: Black, Darla  
; APPLICANT: Schinardiello, Franco  
; APPLICANT: Hilliard, Julia K.  
; TITLE OF INVENTION: Cloning and Amplification of Monkey B  
; NUMBER OF INVENTION: Virus Genes  
; NUMBER OF SEQUENCES: 12  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Cox & Smith Incorporated  
; STREET: 112 East Pecan Street, Suite 2000  
; CITY: San Antonio  
; STATE: Texas  
; COUNTRY: USA  
; ZIP: 78205  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: Patentin Release #1.0, Version #1.25  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/042,747A  
; FILING DATE:  
; CLASSIFICATION: 435  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Haymond, W. Bradley  
; REGISTRATION NUMBER: 35186  
; REFERENCE/DOCKET NUMBER: S-0072.179  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: 210-554-5500  
; TELEFAX: 210-226-8395  
; TELEX: 767609  
; INFORMATION FOR SEQ ID NO: 7:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 2943 base pairs  
; TYPE: nucleic acid  
; STRANDEDNESS: single  
; TOPOLOGY: linear  
; MOLECULE TYPE: Genomic DNA  
; FEATURE:  
; NAME/KEY: CDS  
; LOCATION: 87..2744  
; US-08-042-747A-7

Query Match 8.2%; Score 51.2; DB 1; Length 2943;  
Best Local Similarity 44.5%; Pred. No. 0.064;  
Matches 203; Conservative 0; Mismatches 233; Indels 0; Gaps 0;

QY 23 CGGACCGGGGCGGGAGTCGGGAGACCAAGTGGAGAGAGGCGGGAGATCCAGGTTCCG 82  
Db 208 CCGCCACGCGCCCGCTTCCGAGAGAGAGGTCGCCGAGAGGAGGAGGAGGAGGCTCG 267  
QY 83 CCGCGGAGCGGACTTCTCTGCTGTCGAGCGGCTGCAAGCGGGGTGAGCGGCGAGCGCG 142  
Db 268 CCAACCGCGGCGCGCGCGCCCAAGCGCTGCTGAGCGCGCGCGCCAGCAGCTGCGGGAAG 327  
QY 143 GGGATCTTGAGACCATGGGGCGCGCGGAGCGCCATCTCTGATGCGCTGAGAACCTGA 202

Db 328 ACTGCGGAGATCAAGGCCCGGAGCGGCGAGCGACCTTCTACGTCTGCGCGGCCGA 387  
QY 203 CCGCGGAGAGGCTCAAGATCTACAGCTGAGGCTGCTGTCGCGCGCGCGAGGAGT 262  
Db 368 CCGCGCGCGAGGCTGCTGAGTTCAGAGCAAGCCCGCGCGTCCCGCGCGCGCGCGCGCC 447  
QY 263 ACGGCGCATCCCGCGGGGCGCGCTGCTGTCATGAGCGGCTTGGACCTCAAGCAAGC 322  
Db 448 AGACTACAGCGAGGAGATCGCGTCTTCAAGGAGAACATCGCCCGCTCAAGATTCA 507  
QY 323 TGGTCAGCTTCTACCTGAGACTACGCGCGCGCGCTCAAGCGTAAAGTGGCGCGCA 382  
Db 508 AGCGCACCATGTACTCAAGAGAGCTGACCGCTGCGAGGTTGCGGACCGGTAATCT 567  
QY 383 TGGGCTGCGAGGATGGCGGCGAGCTGCAAGCGGCGCGACGACGAGGCTGAGCGGCC 442  
Db 568 CGCAGTTATGGGATCTTGAAGACCGCGCGCGCGCGCTTCCGAGGAGTATGAGCA 627  
QY 443 CGTCCCGCTCAGCGGCTTTCGCGTCCGACCCACA 478  
Db 628 AGATCAACGCCAAGGGGCTCTCGCGTGCAGCGCCA 663

RESULT 5  
US-08-125-468-1/c  
; Sequence 1, Application US/08125468  
; Patent No. 5589385  
; GENERAL INFORMATION:  
; APPLICANT: Ryan, Michael J.  
; APPLICANT: Lotvin, Jason A.  
; APPLICANT: Strathy, Nancy  
; APPLICANT: Fantihi, Susan E.  
; TITLE OF INVENTION: Cloning of the biosynthetic pathway for  
; TITLE OF INVENTION: chlorotetracycline and tetracycline formation and cosmid  
; NUMBER OF INVENTION: useful therein  
; NUMBER OF SEQUENCES: 1  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: American Cyanamid Company  
; STREET: One Cyanamid Plaza  
; CITY: Wayne  
; STATE: New Jersey  
; COUNTRY: USA  
; ZIP: 07470  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: Patentin Release #1.0, Version #1.25  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/125,468  
; FILING DATE: 22-SEP-1993  
; CLASSIFICATION: 435  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Tseydos, Estelle J  
; REGISTRATION NUMBER: 31,145  
; REFERENCE/DOCKET NUMBER: 31,255-02  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: (201)831-3241  
; TELEFAX: (201)831-3305  
; INFORMATION FOR SEQ ID NO: 1:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 30001 base pairs  
; TYPE: nucleic acid  
; STRANDEDNESS: single  
; TOPOLOGY: linear  
; MOLECULE TYPE: DNA (genomic)  
; US-08-125-468-1

Query Match 8.1%; Score 50.8; DB 1; Length 30001;  
Best Local Similarity 45.6%; Pred. No. 0.083;  
Matches 219; Conservative 0; Mismatches 257; Indels 4; Gaps 1;

QY 3 CGCCTTTGCTGAGGCGCAAGCGACCGGGGCGGGAGTCCGAGACCAAGTGGAGAGAA 62

```

Db 24497 CGCGGCTTCACAGACCGCGATCTGCTCCGCCGAGCGGAGTGGAGACCGGGGATGACCGAC 24438
Qy 63 GCGGGGGAATCCAGGTTCCGCGCCCGGAGACCGGACATTCCTCTGTCGGCGGCTGCACACGGG 122
Db 24437 GTCTGTGTATCCGGGATGCGC---GTGTGTGTCACCGAACCCTGTGGGACCTTGTGGGATGA 24382
Qy 123 GTAGCGCGCGCAGCGCGCCGGGATCTCTGAGGCATGGGGGCGCGCGCGACGCCATCTCT 182
Db 24381 CTGGGACGCGCACACTGTGGCGGGCGCGTCCGGTCTCGGGGCGCGTGCACCCGTTTCAGTTCAC 24322
Qy 183 GGAATGGCGTGGAGAACCTGCAGACCGCGGAGAGCTCAGAAAGATTCAAGCTGAAGTGGTGTTC 242
Db 24321 CGGGTACCGGGGGCGACGTCGGCGCGGAGGGTAATTCACGCGGGAGGAACACTGCGCGGG 24282
Qy 243 GGTGCGCGCTGCGCGGAGGCTACGGGCGCATCCCGCGGGGCGCGGCTGCTGTCTCATGAGACG 302
Db 24261 CCGACTGCTGCGCGGAGACCGACCATGACGCGGCTGGGCTGGGCTGCGCGCGCAGGAGGCG 24202
Qy 303 CTGTGACCTCACCGGACCAAGCTGTGTCAGCTTCACTCGGAGAGACTACGCGGCGCGAGCTCAC 362
Db 24201 GCTGCGCGACGCGGGGGGCGGACCGGACGAGATGCGGAGCTTCTCCCGCGGCTGTGCAC 24142
Qy 363 CGCTAACGTGTCTGCGCGACATGGGCTGTGACGAGATGGCGCGGACGCTGACGAGCGCGCAC 422
Db 24141 GCGCGCTCTGCGCGCGCGGCTTCGAGTTCCGCGCAGAAAGAACTCAAGCGCGCTGTGGAGCAA 24082
Qy 423 GCACGAGGGGTGAGCGCGCCCGCTTCCCTCCCTCAGCCCGCTTTCCTCCCTCACCGACGACG 482
Db 24081 GGGCGCGGCTACGTCAAGCGCTACCAAGTCTTTCGGTGTGTTTACCGCGGTCAACCGGG 24022

RESULT 6
US-08-474-933-1/c
: Sequence 1, Application US/08474933
: Patent No. 5866410
GENERAL INFORMATION:
APPLICANT: Ryan, Michael J.
APPLICANT: Lotvin, Jason A.
APPLICANT: Strathly, Nancy
APPLICANT: Fantlil, Susan E.
TITLE OF INVENTION: Cloning of the biosynthetic pathway for
TITLE OF INVENTION: chlorotetracycline and tetracycline formation and cosmids
NUMBER OF SEQUENCES: 1
CORRESPONDENCE ADDRESS:
ADDRESSEE: American Cyanamid Company
STREET: One Cyanamid Plaza
City: Wayne
STATE: New Jersey
COUNTRY: USA
ZIP: 07470
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentln Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/474,933
FILING DATE:
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/125,468
FILING DATE: 22-SEP-1993
ATTORNEY/AGENT INFORMATION:
NAME: Tsevdos, Estelle J
REGISTRATION NUMBER: 31,145
REFERENCE/DOCKET NUMBER: 31,255-02
TELECOMMUNICATION INFORMATION:
TELEPHONE: (201)831-3241
TELEFAX: (201)831-3305
INFORMATION FOR SEQ ID NO: 1:
SEQUENCE CHARACTERISTICS:

```

```

:      LENGTH: 30001 base pairs
:      TYPE: nucleic acid
:      STRANDEDNESS: single
:      TOPOLOGY: linear
:      MOLECULE TYPE: DNA (genomic)
:      US-08-474-933-1

Query Match
Best Local Similarity 8.1%; Score 50.8; DB 2; Length 30001;
Matches 219; Conservative 0; Mismatches 257; Indels 4; Gaps 1

DB 24497 GGGGGGAGTCCAGATTCCGCCCCGGAGCCGACTTCTCCTGTGTGGCGGCTGCAGCGGG 122
QY 3 CGCCTTTTGTCTGGAGGGGCAACGACGACCGGGCGCGGAGCTGCGGGAGACCAAGTGGAGGAA 62
DB 24497 GCGCGGCTTCCAGAGCCGCGATCTCTCTCTCCAGCCGAGTGAAGCCGGGATGACCGAC 24438
QY 63 GGGGGGAGTCCAGATTCCGCCCCGGAGCCGACTTCTCCTGTGTGGCGGCTGCAGCGGG 122
DB 24437 GTCTGTGTGATCCCGGGATCGGCG-----GTGTGGCACACGAAAGCGCTGGGGCGTTTGCGGAGTA 24382
QY 123 GTGAGCGGGGCGACGGCGCGGGGATCTGTGAGACCATGAGGGCGCGCGCGAGCCGACCACT 182
DB 24381 CTGGGAGCGCACACTGGCGGGCGCGGTCTGGGTCTGGGCGCCTTACCCGGTTTGACTTCAC 24322
QY 183 GGTATGCGCTGGAGAACCTGACCGCCGAGGAGCTCAAGAAGTTTCAAGCTGAAGCTGTCTC 242
DB 24321 CGGGTACCGGGCGCACCTGGCGCGGAGGTGAATTCACCCGAGAAACACTCTCCGGG 24262
QY 243 GGTGGCGCGTGGGGGAGGGCTAACGGGCGCATCCCGCGGGGGGGCGCTGTCTGTCCAGAGCG 302
DB 24261 CGGACTGCTGCCGAGACCGACCACTGACGCGGCTGGCCCTGTGTGCGCGCGGAGAGAGC 24202
QY 303 CTTTGACCTCACCGACCAAGCTGTGCTACCTTACTGTGAGAGCTTACGAGCGCGGAGCTCAC 362
DB 24201 GCTCCGCCACCGGGGGGCGGACCCGACGACGATCCGAGCTTCTCCGCGGGGCTGTGTGAC 24142
QY 363 CGGTACGCTGTGGGGGACATGTGGGCGCTGTGAGAGATGGCCGGGCGAGCTGTGAGGCGCAC 422
DB 24141 GGGCGGCTTCGGCGCGCGGCTGTGAGTGTGGCGCAGAGAACTTCAAGCGGCTGTGTGAGCA 24082
QY 423 GCACGAGGGTGAAGCGCGCCCGTTCCCTCCCTCACCCCGTCTTTCCCTCCACCCACGACAG 482
DB 24081 GGGCGCGGCTACGTACGACGCGCTACGACGATCTGCGCTGTCTTACCGTCAACACGCG 24022

RESULT 7
US-08-476-519-10
:      Sequence 10, Application US/08476519
:      Patent No. 5750876
:      GENERAL INFORMATION:
:      APPLICANT: Barry, Gerard F.
:      APPLICANT: Kishore, Ganesh M.
:      APPLICANT: Krohn, Bradley M.
:      TITLE OF INVENTION: No. 5750876e1 Isoamylase Gene, Compositions
:      TITLE OF INVENTION: Containing It and Methods of Using Isoamylases
:      NUMBER OF SEQUENCES: 11
:      CORRESPONDENCE ADDRESS:
:      ADDRESSEE: Grace L. Bonner, Monsanto Company, BB4F
:      STREET: 700 Chesterfield Parkway No. 5750876th
:      CITY: St. Louis
:      STATE: Missouri
:      COUNTRY: USA
:      ZIP: 63198
:      COMPUTER READABLE FORM:
:      MEDIUM TYPE: Floppy disk
:      COMPUTER: IBM PC compatible
:      OPERATING SYSTEM: PC-DOS/MS-DOS
:      SOFTWARE: PatentIn Release #1.0, Version #1.25
:      CURRENT APPLICATION DATA:
:      FILING DATE: US/08/476,519
:      CLASSIFICATION: 800
:      PRIOR APPLICATION DATA:
:      APPLICATION NUMBER: US 08/281902

```

```

; FILING DATE: 28-JUL-1994
; ATTORNEY/AGENT INFORMATION:
; NAME: Bonner, Grace L.
; REGISTRATION NUMBER: 32,963
; REFERENCE/DOCKET NUMBER: 38-21(13577)A
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (314)537-7286
; TELEFAX: (314)537-6047
; INFORMATION FOR SEQ ID NO: 10:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 2244 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: double
; TOPOLOGY: linear
; MOLECULE TYPE: DNA (genomic)
; FEATURE:
; NAME/KEY: CDS
; LOCATION: 1..2241
;
US-08-476-519-10

```

```

Query Match      8.0%; Score 50; DB 1; Length 2244;
Best Local Similarity 44.9%; Pred. No. 0.11;
Matches 230; Conservative 0; Mismatches 280; Indels 2; Gaps 1;

```

```

QY 15 GAGGCAACGACGACCGGCGGAGTCGGAGACCAAGTGGAGGAAAGCGGGAGTCC 74
DB 1339 GAGTGAACGGCTCTACCGGACGCGCTGCGCAAGAAACAAAGCTCGCGTGGA 1398
QY 75 AGTTCCGCCCCCGAGCCGACCTTCTCTGTCGGCGCTCAGCGGGGTAGCGCGGC 134
DB 1399 ACGGTACACCCCGGACCCCTTGCCACGCGCTTCCGCGCTCAACGACCTGTACGGCGAC 1458
QY 135 AGCGCGCGGAGTCTGG--AGCCATGGGGCGCGCGCGCGACGCGCATCTGATGCGCTG 192
DB 1459 GACGGCGCGAAGCGGTGSCATGTGATCACTTGTGTGTCGCCACGACGGCTTACCCCTC 1518
QY 193 GAGAACCTGACCGCCGAGAGCTCAAGAAATTCAAGCTGAAGCTGTGTGCGTCCGCTG 252
DB 1519 AAGGACCTGTACGCTTCAACGACGACAAACCAACGCGGTGCGTACGGGCGCTCC 1578
QY 253 CGGAGAGGCTACGAGGCGCATCCCGCGGGCGCGCTGTCTCATGAGCGCTTGGACCTC 312
DB 1579 GACGGCGCGAGGAGCAACACTGAGCTGGAACCAAGGGGGGATGTGTGCGCGAGCGCG 1638
QY 313 ACCGACAACTGTGTGCTTCTACCTGTGAGACCTTACGGCGCGGACCTACCGCTAACGTG 372
DB 1639 AAGGCGCGGACGACCGGACCTGAGTGTGCTGATGCTCAAGCGCGCGCTGCGATGATCAC 1698
QY 373 CTGCGGACATGGGCTTGCAGAGATGGCCGCGGACGCTGACAGCGCGCGACGACAGGCT 432
DB 1699 GCGGCGGACGAGGCGCTGCGGACCGACGATTCGGGCAACAAACACTGAACTGGAATTCG 1758
QY 433 GAGCGCGCCCGCTTCCCTCCACCCGCTTTCCTCCACCGACACGACGCTTACCC 492
DB 1759 GCGGCGCAACTGGCTGTATGTAGAGCCGAGCGCGCTGAGAGCGGACCAAGACCTTACCC 1818
QY 493 GCGGCGCTTCCGCTTCTGTCTTCTCTCTACCC 524
DB 1819 AAGCGCTGATCGGCTTCCGCAAGGCGCACCC 1850

```

```

RESULT 8
PCT-US95-09323-10
; Sequence 10, Application PC/TUS9509323
; GENERAL INFORMATION:
; APPLICANT:
; TITLE OF INVENTION: Novel Isoamylase Gene, Compositions
; TITLE OF INVENTION: Containing It and Methods of Using Isoamylases
; NUMBER OF SEQUENCES: 11
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS

```

```

; SOFTWARE: PatentIn Release #1.0, Version #1.30 (EPO)
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: PCT/US95/09323
; FILING DATE:
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/281902
; FILING DATE: 28-JUL-1994
; INFORMATION FOR SEQ ID NO: 10:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 2244 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: double
; TOPOLOGY: linear
; MOLECULE TYPE: DNA (genomic)
; FEATURE:
; NAME/KEY: CDS
; LOCATION: 1..2241
;
PCT-US95-09323-10

```

```

Query Match      8.0%; Score 50; DB 5; Length 2244;
Best Local Similarity 44.9%; Pred. No. 0.11;
Matches 230; Conservative 0; Mismatches 280; Indels 2; Gaps 1;

```

```

QY 15 GAGGCAACGACGACCGGCGGAGTCGGAGACCAAGTGGAGGAAAGCGGGAGTCC 74
DB 1339 GAGTGAACGGCTCTACCGGACGCGCTGCGCAAGAAACAAAGCTCGCGTGGA 1398
QY 75 AGTTCCGCCCCCGAGCCGACCTTCTCTGTCGGCGCTCAGCGGGGTAGCGCGGC 134
DB 1399 ACGGTACACCCCGGACCCCTTGCCACGCGCTTCCGCGCTCAACGACCTGTACGGCGAC 1458
QY 135 AGCGCGCGGAGTCTGG--AGCCATGGGGCGCGCGCGCGACGCGCATCTGATGCGCTG 192
DB 1459 GACGGCGCGAAGCGGTGSCATGTGATCACTTGTGTGTCGCCACGACGGCTTACCCCTC 1518
QY 193 GAGAACCTGACCGCCGAGAGCTCAAGAAATTCAAGCTGAAGCTGTGTGCGTCCGCTG 252
DB 1519 AAGGACCTGTACGCTTCAACGACGACAAACCAACGCGGTGCGTACGGGCGCTCC 1578
QY 253 CGGAGAGGCTACGAGGCGCATCCCGCGGGCGCGCTGTCTCATGAGCGCTTGGACCTC 312
DB 1579 GACGGCGCGAGGAGCAACACTGAGCTGGAACCAAGGGGGGATGTGTGCGCGAGCGCG 1638
QY 313 ACCGACAACTGTGTGCTTCTACCTGTGAGACCTTACGGCGCGGACCTACCGCTAACGTG 372
DB 1639 AAGGCGCGGACGACCGGACCTGAGTGTGCTGATGCTCAAGCGCGCGCTGCGATGATCAC 1698
QY 373 CTGCGGACATGGGCTTGCAGAGATGGCCGCGGACGCTGACAGCGCGCGACGACAGGCT 432
DB 1699 GCGGCGGACGAGGCGCTGCGGACCGACGATTCGGGCAACAAACACTGAACTGGAATTCG 1758
QY 433 GAGCGCGCCCGCTTCCCTCCACCCGCTTTCCTCCACCGACACGACGCTTACCC 492
DB 1759 GCGGCGCAACTGGCTGTATGTAGAGCCGAGCGCGCTGAGAGCGGACCAAGACCTTACCC 1818
QY 493 GCGGCGCTTCCGCTTCTGTCTTCTCTCTACCC 524
DB 1819 AAGCGCTGATCGGCTTCCGCAAGGCGCACCC 1850

```

```

RESULT 9
US-08-476-519-1
; Sequence 1, Application US/08476519
; Patent No. 5750876
; GENERAL INFORMATION:
; APPLICANT: Barry, Gerard F.
; APPLICANT: Kishore, Ganesh M.
; APPLICANT: Krohn, Bradley M.
; TITLE OF INVENTION: No. 3750876el Isoamylase Gene, Compositions
; TITLE OF INVENTION: Containing It and Methods of Using Isoamylases
; NUMBER OF SEQUENCES: 11
; CORRESPONDENCE ADDRESS:

```

ADDRESS: Grace L. Bonner, Monsanto Company, Bldg. 100  
STREET: 700 Chesterfield Parkway No. 5750876th  
CITY: St. Louis  
STATE: Missouri  
COUNTRY: USA  
ZIP: 63198  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: PatentIn Release #1.0, Version #1.25  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/476,519  
FILING DATE:  
CLASSIFICATION: 800  
PRIORITY APPLICATION DATA:  
APPLICATION NUMBER: US 08/281902  
FILING DATE: 28-JUL-1994  
ATTORNEY/AGENT INFORMATION:  
NAME: Bonner, Grace L.  
REGISTRATION NUMBER: 32,963  
REFERENCE/DOCKET NUMBER: 38-21(13577)A  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (314)537-7286  
TELEFAX: (314)537-6047  
INFORMATION FOR SEQ. ID NO.: 1:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 2334 base pairs  
TYPE: nucleic acid  
STRANDEDNESS: double  
TOPOLOGY: Linear  
MOLECULE TYPE: DNA (genomic)  
IS-08-476-519-1

```
US-08-459-586-15
; Sequence 15, Application US/08459586
; Patent No. 5720957
; GENERAL INFORMATION:
; APPLICANT: Jones, Thomas R.
; APPLICANT: Campbell, Ann E.
; TITLE OF INVENTION: Identification of a Human
; TITLE OF INVENTION: Cytomegalovirus Gene Region Involved in Down-Regulation of
; TITLE OF INVENTION: MHC Class I Heavy Chain Expression
; NUMBER OF SEQUENCES: 20
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: American Cyanamid Company
; STREET: One Campus Drive
; CITY: Parsippany
; STATE: New Jersey
; COUNTRY: United States
; ZIP: 07054
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patentin Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/459,586
; FILING DATE: 02-JUN-1995
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: Barnhard, Elizabeth M.
; REGISTRATION NUMBER: 31,088
; REFERENCE/DOCKET NUMBER: 32,425-01
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 201-683-2158
; TELEFAX: 201-683-4117
; INFORMATION FOR SEQ ID NO: 15:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1765 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: DNA (genomic)
US-08-459-586-15

Query Match          7.9%; Score 49.4; DB 1; Length 1765;
Best Local Similarity 44.9%; Pred. No. 0.14;
Matches 188; Conservative 0; Mismatches 231; Indels 0; Gaps 0;

QY 23 CGGACCGGGGCGGGAGTCGGGAGACGAGTGGAGAGGCGGGGAGTCCAGGTTCCG 82
DB 104 CGTACGGGCGCGCGGGCGGGGCGGATTTGCGCTTACGCGGAGCGCGCGGTGG 163
QY 83 CCCCAGGCGCGAGTTCCTCTGTCGGGGGCGTCGACGCGGGGTGAGCGGGGAGAGCGCG 142
DB 164 TCGGGGCGGAGCGGAGACGACGACGCGGCGGAGGATGATGGTGGTTGACAGGATGG 223
QY 143 GGGATCTTGAGACCATGGGGCGCGCGCGACGACCATCTGATGGCGCTGAGAACTGA 202
DB 224 CGGGTCCGGGGCGGTCGCGCGCGGCGGCGCATGCTGGTGGTGGCTGAGCGAGC 283
QY 203 CCGGCGGAGGCTCAGAGATTCAAGCTGAGCTGATGCTGTCGTCGCGCGCGGAGGCT 262
DB 264 TGGGCGCGGTTGTGGGTCTGCTCCGCTGGAAGGGACGCTGATCCCGCTGGCGGGGAGC 343
QY 263 ACGGGCGCATCCGCGGGGCGCGCTGCTGTCATGAGACGCTTGAACCTGACCGAAGC 322
DB 344 TGTGCGACTTTCTGCGCGGGGCGTGTGCTGGCGCGCTGGGAGCCGAGTGGGCGC 403
QY 323 TGTGCTACTTCTACCTGAGAGCTACGCGCGGCGGCTCACCGCTTAACGTCGCGCGACA 382
DB 404 CCGCGCGCGAGCGCGCGCGGCGGCTGCTGCTGATGAGCGGCGGAGAGTGGAGCGGC 463
QY 383 TGGGCTGTCAGAGATGCGCGGGGAGCTGACGAGCGCGGCGACGACGAGGTTGAGCGGCC 441
DB 464 CGCGCTGGGAGCGCGCTGACACCGCGCGCGCGCGCTGTGGCGCGCGAGCGCGAC 522
```

```
RESULT 12
US-08-282-696-15
; Sequence 15, Application US/08282696
; Patent No. 5846806
; GENERAL INFORMATION:
; APPLICANT: Jones, Thomas R.
; APPLICANT: Campbell, Ann E.
; TITLE OF INVENTION: Identification of a Human
; TITLE OF INVENTION: Cytomegalovirus Gene Region Involved in Down-Regulation of
; TITLE OF INVENTION: MHC Class I Heavy Chain Expression
; NUMBER OF SEQUENCES: 20
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: American Cyanamid Company
; STREET: One Cyanamid Plaza
; CITY: Wayne
; STATE: New Jersey
; COUNTRY: United States
; ZIP: 07470-8426
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; OPERATING SYSTEM: IBM PC compatible
; SOFTWARE: Patentin Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/282,696
; FILING DATE: 29-JUL-1994
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: Barnhard, Elizabeth M.
; REGISTRATION NUMBER: 31,088
; REFERENCE/DOCKET NUMBER: 32,425-00
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 201-831-3246
; TELEFAX: 201-831-3305
; INFORMATION FOR SEQ ID NO: 15:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1765 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: DNA (genomic)
US-08-282-696-15

Query Match          7.9%; Score 49.4; DB 2; Length 1765;
Best Local Similarity 44.9%; Pred. No. 0.14;
Matches 188; Conservative 0; Mismatches 231; Indels 0; Gaps 0;

QY 23 CGGACCGGGGCGGGAGTCGGGAGACGAGTGGGAGAGGCGGGGAGTCCAGGTTCCG 82
DB 104 CGTACGGGCGCGCGGGCGGGGCGGATTTGCGCTTACGCGGAGCGCGCGGTGG 163
QY 83 CCCCAGGCGCGAGTTCCTCTGTCGGGGGCGTCGACGCGGGGTGAGCGGGGAGAGCGCG 142
DB 164 TCGGGGCGGAGCGGCGGCGGCGGCGGCGGCGGCGGATGATGGTGGTTGACAGGATGG 223
QY 143 GGGATCTTGAGACCATGGGGCGCGCGCGACGACCATCTGATGGCGCTGAGAACTGA 202
DB 224 CGGGTCCGGGGCGGTCGCGCGCGGCGGCGCATGCTGGTGGTGGCTGAGCGAGC 283
QY 203 CCGGCGGAGGCTCAGAGATTCAAGCTGAGCTGATGCTGTCGTCGTCGCGCGCGGAGGCT 262
DB 264 TGGGCGCGGTTGTGGGTCTGCTCCGCTGGAAGGGACGCTGATCCCGCTGGCGGGGAGC 343
QY 263 ACGGGCGCATCCGCGGGGCGCGCTGCTGTCATGAGACGCTTGAACCTGACCGAAGC 322
DB 344 TGTGCGACTTTCTGCGCGGGGCGTGTGCTGGCGCGCTGGGAGCCGAGTGGGCGC 403
QY 323 TGTGCTACTTCTACCTGAGAGCTACGCGCGGCGGCTCACCGCTTAACGTCGCGCGACA 382
DB 404 CCGCGCGCGAGCGCGCGGCGGCGTGTGCTGCTGATGAGCGGCGGAGAGTGGAGCGGC 463
```



Db 14384 TGTGCACTTTCTGCGCGGCGGCTGCTG6GCGCCTG6GCGTGGAGCGCAGTCGCGC 14443  
QY 323 TGTGACGCTTACCTGAGACCTAGCGGCGCGAGCTGACCGCTACGCTGCTGCGGACA 382  
Db 14444 CCGCCCGCCGAGCGCGCGGCGGCTGCTGCGCGAGCTGGACCGCGAGCATGGAGCGGC 14503  
QY 383 TGGGCGCTGAGAGATGCGCGCGAGCTGCGAGCGCGCACGACACAGGCTGAGCGCGCC 441  
Db 14504 CGCGCTGGAGACGCTGACCTGCACCGCGCGCGCTGTGGGCGCGGAGCGGCAC 14562

## RESULT 15

US-08-758-662-4  
; Sequence 4, Application US/08758662  
; Patent No. 6114150  
; GENERAL INFORMATION:  
; APPLICANT: Weisman, Sherman  
; APPLICANT: Baskaran, Namadev  
; TITLE OF INVENTION: AMPLIFICATION OF NUCLEIC ACIDS  
; NUMBER OF SEQUENCES: 8  
; CORRESPONDENCE ADDRESSES:  
; ADDRESSEE: SEED and BERRY LLP  
; STREET: 701 Fifth Avenue, 6300 Columbia Center  
; City: Seattle  
; STATE: Washington  
; COUNTRY: USA  
; ZIP: 98104  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: Patent Release #1.0, Version #1.30  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/758,662  
; FILING DATE: 29-NOV-1996  
; CLASSIFICATION: 435  
; ATTORNEY/AGENT INFORMATION:  
; NAME: No. 6114150Lendburg Ph.D., Carol  
; REGISTRATION NUMBER: 39,317  
; REFERENCE/DOCKET NUMBER: 390036.402C1  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: (206) 622-4900  
; TELEFAX: (206) 682-6031  
; INFORMATION FOR SEQ ID NO: 4:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 530 base pairs  
; TYPE: nucleic acid  
; STRANDEDNESS: single  
; TOPOLOGY: linear  
; US-08-758-662-4

Query Match 7.7%; Score 48; DB 3; Length 530;

Best Local Similarity 55.4%; Pred. No. 0.25; Mismatches 75; Indels 0; Gaps 0;

Matches 93; Conservative 0; Mismatches 75; Indels 0; Gaps 0;  
QY 23 CGAGCCGGGGCGGAGATCGGAGACCAAGTGGAGGAGGAGGAGTCCAGGTTCCG 82  
Db 310 CGGCGGCGGCGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 369  
QY 83 CCGCGGAGCGGAGCTTCCTGTGCGGCGGCTGACGAGGCGGCTGACGAGGCGGCGGCGG 142  
Db 370 CCGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 429  
QY 143 GGAATCCTGTGAGCATTGGGGCGCGCGGCGGAGCGCATTCCTGATGCGC 190  
Db 430 GGGGCGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 477

Search completed: December 22, 2002, 06:48:32  
Job time : 142 secs





GenCore version 5.1.3  
Copyright (c) 1993 - 2002 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: December 22, 2002, 05:23:21 | Search time 1992 Seconds  
(without alignments)  
5089.546 Million cell updates/sec

Title: US-09-691-763b-4

Perfect score: 626  
Sequence: 1 aggcgccttgcgtggaagcgc.....ggaagggaagggatcactt 626

Scoring table:  
IDENTITY:NUC  
Gapop 10.0 , Gapext 1.0

Searched: 16154066 seqs, 8097743376 residues

Total number of hits satisfying chosen parameters: 32308132

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

EST:\*  
1: em\_estba:\*  
2: em\_esthum:\*  
3: em\_estlin:\*  
4: em\_estmu:\*  
5: em\_estov:\*  
6: em\_estopl:\*  
7: em\_estro:\*  
8: em\_hic:\*  
9: gq\_estl:\*  
10: gq\_est2:\*  
11: gq\_hic:\*  
12: gq\_est3:\*  
13: gq\_est4:\*  
14: gq\_est5:\*  
15: em\_estfun:\*  
16: em\_estom:\*  
17: gq\_gss:\*  
18: em\_gss\_hum:\*  
19: em\_gss\_iny:\*  
20: em\_gss\_pln:\*  
21: em\_gss\_vit:\*  
22: em\_gss\_fun:\*  
23: em\_gss\_man:\*  
24: em\_gss\_mus:\*  
25: em\_gss\_other:\*  
26: em\_gss\_pro:\*  
27: em\_gss\_rtd:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	424.4	67.8	866	12	BG684386
2	340.4	54.4	741	13	BM051141
3	338	54.0	767	14	BM910918
4	337.2	53.9	1079	13	BM456838
5	321	51.3	744	12	BG255521
6	320	51.1	320	13	BT756931

7	320	51.1	930	13	BM459678
8	317	50.6	830	12	BF238010
9	316	50.5	792	13	BM459665
10	315	50.3	804	14	BM926438
11	314	50.2	750	12	BG764161
12	314	50.2	856	12	BG767422
13	312.4	49.9	748	12	BG337806
14	301.2	48.1	1294	14	BM918855
15	296	47.3	446	14	BM726665
16	270.4	43.2	634	14	AT991236
17	270	43.1	687	12	BE909218
18	259	41.4	888	12	BE906286
19	258.4	41.3	678	9	AT148558
20	248	39.6	676	14	BM972785
21	247	39.5	675	12	BE908204
22	244.4	39.0	615	9	AT346818
23	241.8	38.6	267	13	BT025158
24	226.6	36.2	308	14	H16108
25	224.4	35.8	262	14	W17108
26	222.8	35.6	560	13	BM364268
27	221.4	35.4	1093	14	BM920038
28	220.4	35.2	830	10	BE560228
29	218.6	34.9	723	10	BE560247
30	218.4	34.9	555	10	AV664060
31	214	34.2	446	9	AA442018
32	211.6	33.8	506	13	BM105270
33	211.2	33.7	404	10	AM479270
34	206	32.9	448	13	BT339834
35	206	32.9	633	14	BM681310
36	205.2	32.8	638	14	BM998354
37	204	32.6	417	13	BM363061
38	204	32.6	427	13	BM365021
39	204	32.6	469	13	BM363361
40	204	32.6	565	13	BM362328
41	200.4	32.0	352	13	BM365886
42	200.4	32.0	367	13	BM365745
43	200.4	32.0	412	13	BM365704
44	200.4	32.0	456	13	BM365720
45	200.4	32.0	473	13	BM255820

#### ALIGNMENTS

RESULT 1  
LOCUS BG684386 866 bp mRNA linear EST 01-MAY-2001  
DEFINITION 602636137F1 NIH\_MGC\_48 Homo sapiens cDNA clone IMAGE:4763985 5',  
mRNA sequence.  
ACCESSION BG684386  
VERSION BG684386.1 GI:13915783  
KEYWORDS EST.  
SOURCE human.  
ORGANISM Homo sapiens  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.  
REFERENCE 1 (bases 1 to 866)  
AUTHORS NIH-MGC http://mgs.nci.nih.gov/.  
TITLES National Institutes of Health, Mammalian Gene Collection (MGC)  
JOURNLS Unpublished (1999)  
COMMENT Contact: Robert Strausberg, Ph.D.  
Email: cgabs-remail.nih.gov  
Tissue Procurement: Louis M. Staudt, M.D., Ph.D.  
CDNA Library Preparation: Ling Hong/Rudin Laboratory  
CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LIML)  
DNA Sequencing by: Incyte Genomics, Inc.  
Clone distribution: MGC clone distribution information can be  
found through the I.M.A.G.E. Consortium/LLNL at:  
http://image.llnl.gov  
Plate: L1CM1620 row: i column: 10  
High quality sequence stop: 862.  
Location/Qualifiers L. 866

FEATURES  
source

/organism="Homo sapiens"  
/db\_xref="taxon:9606"  
/clone="IMAGE:4763985"  
/clone\_1lb="NIH\_MGC\_48"  
/tissue\_type="Primary B-cells from tonsils (cell line)"  
/lab\_host="DH10B (phage-resistant)"  
/note="Organ: B-cells; Vector: pORF7; Site\_1: XhoI;  
Site\_2: EcoRI; cDNA made by oligo-dT priming.  
Directionally cloned into EcoRI/XhoI sites using the  
following 5' adaptor: GGCACGAG(G). Size selected >500bp  
for average insert size 1.8kb. Library constructed by Ling  
Hong in the laboratory of Gerald M. Rubin (University of  
California, Berkeley) using ZAP-cDNA synthesis kit  
(Stratagene) and Superscript II RT (Life Technologies).  
Note: this is a NIH\_MGC library."

BASE COUNT 154 a 279 c 310 g 123 t  
ORIGIN

Query Match 67.8%; Score 424.4; DB 12; Length 866;  
Best Local Similarity 99.8%; Pred. No. 3.8e-69;  
Matches 425; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

OY 6 CTTTGGTGGAGGCAAGGACCGGGGGGAGTGGGAGACCAAGTGGAGGAGGAGGC 65  
DB 132 CTTTGGTGGAGGCAAGGACCGGGGGGAGTGGGAGACCAAGTGGAGGAGGAGGC 211  
OY 66 GGGAGTCCAGGTTCCGCCCGGAGCCGACTTCTCTGGTGGGCGCTGCAGCGGGGTG 125  
DB 212 GGGAGTCCAGGTTCCGCCCGGAGCCGACTTCTCTGGTGGGCGCTGCAGCGGGGTG 271  
OY 126 AGCGGGGCGAGCGGGGGGATCTGAGACCATGGGGCGGGCGGACCGCATCTGGA 185  
DB 272 AGCGGGGCGAGCGGGGGGATCTGAGACCATGGGGCGGGCGGACCGCATCTGGA 331  
OY 186 TGGCTGGAGAACCTTACCGCCGAGAGCTCAGAGTTCAGCTGAGCTGCTGCT 245  
DB 332 TGGCTGGAGAACCTTACCGCCGAGAGCTCAGAGTTCAGCTGAGCTGCTGCT 391  
OY 246 GCCCTGCGGAGGGGTACGGGGCGATCCCGGGGGCGCTCTGCTCATGAGACGCTT 305  
DB 392 GCCCTGCGGAGGGGTACGGGGCGATCCCGGGGGCGCTCTGCTCATGAGACGCTT 451  
OY 306 GGACCTCAGCAGAACCTGTGACTTCTACCTGAGACCTTACGGCGCCGAGCTACCGC 365  
DB 452 GGACCTCAGCAGAACCTGTGACTTCTACCTGAGACCTTACGGCGCCGAGCTACCGC 511  
OY 366 TTAAGCTCTCGCGACATGGGCTTCAGAGAGATGGCGGGCAGCTTCAGGCGCAGCA 425  
DB 512 TTAAGCTCTCGCGACATGGGCTTCAGAGAGATGGCGGGCAGCTTCAGGCGCAGCA 571  
OY 426 CCAGG 431  
DB 572 CCAGG 577

RESULT 2  
LOCUS BM051141 741 bp mRNA linear EST 07-NOV-2001  
DEFINITION 603634063F1 NIH\_MGC\_43 Homo sapiens cDNA clone IMAGE:5424590 5',  
mRNA sequence.  
ACCESSION BM051141  
VERSION BM051141.1 GI:16780408  
KEYWORDS EST.  
SOURCE human.  
ORGANISM Homo sapiens

REFERENCE 1 (bases 1 to 741)  
AUTHORS NIH-MGC http://mgc.nci.nih.gov/.  
TITLE National Institutes of Health, Mammalian Gene Collection (MGC)  
JOURNAL Unpublished (1999)  
COMMENT Contact: Robert Strausberg, Ph.D.  
Email: cgapbs-r@mail.nih.gov

Tissue Procurement: ATCC  
cDNA Library Preparation: Ling Hong/Rubin Laboratory  
DNA Sequencing by: The I.M.A.G.E. Consortium (LNLN)  
DNA Sequencing by: InCyte Genomics, Inc.  
Clone distribution: MGC clone distribution information can be  
found through the I.M.A.G.E. Consortium/LNLN at:  
http://image.lnl.gov  
Plate: LNCM1885 row: n column: 15  
High quality sequence stop: 643.  
Location/Qualifiers

FEATURES  
source  
1. 741  
/organism="Homo sapiens"  
/db\_xref="taxon:9606"  
/clone="IMAGE:5424590"  
/clone\_1lb="NIH\_MGC\_43"  
/tissue\_type="normal pigmented retinal epithelium"  
/lab\_host="DH10B (phage-resistant)"  
/note="Organ: eye; Vector: pORF7; Site\_1: XhoI; Site\_2:  
EcoRI; cDNA made by oligo-dT priming. Directionally  
cloned into EcoRI/XhoI sites using the following 5'  
adaptor: GGCACGAG(G). Library constructed by Ling Hong  
in the laboratory of Gerald M. Rubin (University of  
California, Berkeley) using ZAP-cDNA synthesis kit  
(Stratagene) and Superscript II RT (Life Technologies).  
Note: this is a NIH\_MGC library."

BASE COUNT 139 a 243 c 237 g 122 t  
ORIGIN

Query Match 54.4%; Score 340.4; DB 13; Length 741;  
Best Local Similarity 99.7%; Pred. No. 1.3e-53;  
Matches 341; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

OY 90 GCCGACTTCCTCTGTCGGCGGCTGACAGCGGCTGAGCGGCGACCGCGGGAGNC 149  
DB 1 GCCGACTTCCTCTGTCGGCGGCTGACAGCGGCTGAGCGGCGACCGCGGGAGTC 60  
OY 150 TGGAGCATGGGGGCGCGCGGACGACCTCTGATCGCTGAGAACCTGACCGCA 209  
DB 61 TGGAGCATGGGGGCGCGCGGACGACCTCTGATCGCTGAGAACCTGACCGCA 120  
OY 210 GAGAGCTCAAGAAAGTCAAGCTGAGAGCTGTCGCTGCGGCGGAGGCGTACGGGCG 269  
DB 121 GAGAGCTCAAGAAAGTCAAGCTGAGAGCTGTCGCTGCGGCGGAGGCGTACGGGCG 180  
OY 270 CATCCCGGGGGCGCGCTGCTGTCATGAGACCTTGGACCTCAGCAGAACCTGATGAG 329  
DB 181 CATCCCGGGGGCGCGCTGCTGTCATGAGACCTTGGACCTCAGCAGAACCTGATGAG 240  
OY 330 CTTTACCTGGAGACTTACGCGCGGAGCTCACCGCTTACGCTGAGGAGACATGGGCT 389  
DB 241 CTTTACCTGGAGACTTACGCGCGGAGCTCACCGCTTACGCTGAGGAGACATGGGCT 300  
OY 390 GCAGAGATGGCGCGGAGCTGCAAGGCGGCGCAGCCAGG 431  
DB 301 GCAGAGATGGCGCGGAGCTGCAAGGCGGCGCAGCCAGG 342

RESULT 3  
LOCUS BM10918 767 bp mRNA linear EST 12-MAR-2002  
DEFINITION 5', mRNA sequence.  
ACCESSION BM10918  
VERSION BM10918.1 GI:19361297  
KEYWORDS EST.  
SOURCE human.  
ORGANISM Homo sapiens

REFERENCE 1 (bases 1 to 767)  
AUTHORS NIH-MGC http://mgc.nci.nih.gov/.  
TITLE National Institutes of Health, Mammalian Gene Collection (MGC)  
JOURNAL Unpublished (1999)

## COMMENT

Contact: Robert Strausberg, Ph.D.  
Email: cgaabs-remail.nih.gov  
Tissue Procurement: ATCC  
cDNA Library Preparation: Rubin Laboratory  
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)  
DNA Sequencing by: Agencourt Bioscience Corporation  
Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LLNL at: <http://image.llnl.gov>  
Plate: L1CM1949 row: 1 column: 18  
High quality sequence stop: 641.

## FEATURES

## source

1..767  
/organism="Homo sapiens"  
/db\_xref="taxon:9606"  
/clone="IMAGE:5454497"  
/clone\_lib="NIH\_MGC\_98"  
/tissue\_type="astrocytoma grade IV, cell line"  
/note="Organ: brain; Vector: pOPB7; Site\_1: XhoI; Site\_2: EcoRI; cDNA made by oligo-3T priming. Directionally cloned into EcoRI/XhoI sites using the following 5' adaptor: GGCACGAG(G). Library constructed by Ling Hong in the laboratory of Gerald M. Rubin (University of California, Berkeley) using ZAP-cDNA synthesis kit (Stratagene) and Superscript II RT (Life Technologies). Note: this is a NIH\_MGC Library."

BASE COUNT 163 a 241 c 237 g 124 t 2 others  
ORIGIN

Query Match 54.0%; Score 338; DB 14; Length 767;  
Best Local Similarity 100.0%; Pred. No. 3.0e-53;  
Matches 338; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 94 ACTTCTCTGCTGGGCGGCTGCAGCGGGGTAGAGCGCGCGCGGGGATCTTGA 153  
Db 1 ACTTCTCTGCTGGGCGGCTGCAGCGGGGTAGAGCGCGCGGGGATCTTGA 60  
QY 154 GCCATGGGGCGCGCGCGGACGCCATCTTGATGCGCTGGAGAACTGACCGCGAGAG 213  
Db 61 GCCATGGGGCGCGCGCGGACGCCATCTTGATGCGCTGGAGAACTGACCGCGAGAG 120  
QY 214 CTCAGAAAGTTCAAGCTGAAGCTGCTGTGGTGGCGCGCGCGGCTTACCGGGCATC 273  
Db 121 CTCAGAAAGTTCAAGCTGAAGCTGCTGTGGTGGCGCGCGGCTTACCGGGCATC 180  
QY 274 CGCGGGCGCGCGCTGCTGCTCATGAGAGCGCTTGACCTACCGAAGCTGTACACTTC 333  
Db 181 CGCGGGCGCGCGCTGCTGCTCATGAGAGCGCTTGACCTACCGAAGCTGTACACTTC 240  
QY 334 TACCTGAGACCTAGCGGCGCGAGCTACCGGCTAAGCTGCTGCGCGAGATGGGCTGAG 393  
Db 241 TACCTGAGACCTAGCGGCGCGAGCTACCGGCTAAGCTGCTGCGCGAGATGGGCTGAG 300  
QY 394 GAGATGGCGGGGAGCTGACAGCGGCGCGACACGACGAGG 431  
Db 301 GAGATGGCGGGGAGCTGACAGCGGCGCGACGACGAGG 338

RESULT 4  
BM56838 1079 bp mRNA linear EST 05-FEB-2002  
LOCUS BM56838  
DEFINITION AGENCOURT\_6404153 NIH\_MGC\_92 Homo sapiens cDNA clone IMAGE:5583760  
5', mRNA sequence.  
ACCESSION BM56838  
VERSION BM56838.1 GI:18505878  
KEYWORDS EST.  
SOURCE human.  
ORGANISM Homo sapiens  
Eukaryota; Metazoa; Chordata; Cranialta; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.  
REFERENCE 1 (bases 1 to 1079)  
AUTHORS NIH-MGC <http://mgc.nci.nih.gov/>.

TITLE  
JOURNAL  
COMMENT

National Institutes of Health, Mammalian Gene Collection (MGC)  
Unpublished (1999)  
Contact: Robert Strausberg, Ph.D.  
Email: cgaabs-remail.nih.gov  
Tissue Procurement: ATCC  
cDNA Library Preparation: Life Technologies, Inc.  
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)  
DNA Sequencing by: Agencourt Bioscience Corporation  
Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LLNL at: <http://image.llnl.gov>  
Plate: L1AM12347 row: 1 column: 17  
High quality sequence stop: 682.

## FEATURES

## source

1..1079  
/organism="Homo sapiens"  
/db\_xref="taxon:9606"  
/clone="IMAGE:5583760"  
/clone\_lib="NIH\_MGC\_92"  
/tissue\_type="embryonal carcinoma, cell line"  
/lab\_host="DH10B (phage-resistant)";  
/note="Organ: testis; Vector: pCMV-SPORT6; Site\_1: NotI; Site\_2: SalI; Cloned unidirectionally; oligo-3T primed. Average insert size 2.5 Kb. Library enriched for full-length clones and constructed by Life Technologies. Note: this is a NIH\_MGC Library."

BASE COUNT 246 a 360 c 297 g 171 t 5 others  
ORIGIN

Query Match 53.9%; Score 337.2; DB 13; Length 1079;  
Best Local Similarity 99.1%; Pred. No. 4.9e-53;  
Matches 339; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 95 CTTCCTCTGCTGGGCGGCTGCAGCGGGGTAGAGCGCGCGCGGGGATCTTGA 154  
Db 1 CTTCCTCTGCTGGGCGGCTGCAGCGGGGTAGAGCGCGCGGGGATCTTGA 60  
QY 155 CCATGGGGCGCGCGCGGACGCCATCTTGATGCGCTGGAGAACTGACCGCGAGAGC 214  
Db 61 CCATGGGGCGCGCGCGGACGCCATCTTGATGCGCTGGAGAACTGACCGCGAGAGC 120  
QY 215 TCAAGAAAGTTCAAGCTGAAGCTGCTGTGGTGGCGCGCGGCTTACCGGGCATC 274  
Db 121 TCAAGAAAGTTCAAGCTGAAGCTGCTGTGGTGGCGCGCGGCTTACCGGGCATC 180  
QY 275 CGCGGGCGCGCGCTGCTGCTCATGAGAGCGCTTGAGCTACCGAAGCTGTACACTTC 334  
Db 181 CGCGGGCGCGCGCTGCTGCTCATGAGAGCGCTTGAGCTACCGAAGCTGTACACTTC 240  
QY 335 ACCTGGAGACCTAGCGGCGCGAGCTACCGGCTAAGCTGCTGCGCGAGATGGGCTGAG 394  
Db 241 ACCTGGAGACCTAGCGGCGCGAGCTACCGGCTAAGCTGCTGCGCGAGATGGGCTGAG 300  
QY 395 AGATGGCGGGGAGCTGACAGCGGCGCGACACGAGGTGAGC 436  
Db 301 AGATGGCGGGGAGCTGACAGCGGCGCGACACGAGGTGAGC 342

RESULT 5  
BG255521 744 bp mRNA linear EST 13-FEB-2001  
LOCUS BG255521  
DEFINITION 602367671F1 NIH\_MGC\_91 Homo sapiens cDNA clone IMAGE:4476088 5',  
mRNA sequence.  
ACCESSION BG255521  
VERSION BG255521.1 GI:12765259  
KEYWORDS EST.  
SOURCE human.  
ORGANISM Homo sapiens  
Eukaryota; Metazoa; Chordata; Cranialta; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.  
REFERENCE 1 (bases 1 to 744)  
AUTHORS NIH-MGC <http://mgc.nci.nih.gov/>.  
TITLE National Institutes of Health, Mammalian Gene Collection (MGC)



TITLE National Institutes of Health, Mammalian Gene Collection (MGC)  
JOURNAL Unpublished (1999)  
COMMENT Contact: Robert Strausberg, Ph.D.  
Email: cgaabs-r@mail.nih.gov  
Tissue Procurement: ATCC

CDNA Library Preparation: Life Technologies, Inc.  
CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)  
DNA Sequencing by: Agencourt Bioscience Corporation  
Clone distribution: MGC clone distribution information can be  
found through the I.M.A.G.E. Consortium/LLNL at:  
<http://image.llnl.gov>  
Plate: LLM12220 row: j column: 05  
High quality sequence stop: 670.

## FEATURES

## source

1. 930  
Location/Qualifiers

/organism="Homo sapiens"  
/db\_xref="taxon:9606"  
/clone="IMAGE:5534308"  
/clone\_lib="NIH\_MGC\_71"  
/tissue\_type="telomerosarcoma"  
/lab\_host="DH10B (phage-resistant)"  
/note="Organ: uterus; Vector: pCMV-Sport6; Site\_1: NotI;  
Site\_2: SalI; Cloned unidirectionally. Primer: Oligo dT.  
Average insert size 2.1 kb."  
BASE COUNT 182 a 310 c 263 g 172 t 3 others  
ORIGIN

Query Match 51.1%; Score 320; DB 13; Length 930;

Best Local Similarity 100.0%; Pred. No. 7.5e-50;

Matches 320; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 112 GCTGACGGGGGTAGAGCGGCGCGGGGATCCTGAGGACATGGGCGCGCGC 171  
DB 1 GCTGACGGGGGTAGAGCGGCGCGGGGATCCTGAGGACATGGGCGCGCGC 60  
QY 172 GACGCGCATCTGATGTGGCTGAGAACCTGACCGCGGAGGCTCAAGATTCAAGCTG 231  
DB 61 GACGCGCATCTGATGTGGCTGAGAACCTGACCGCGGAGGCTCAAGATTCAAGCTG 120  
QY 232 AAGCTGCTGTGCTGCGCGCGGAGGAGGCTACCGCGGCGCGCTGCTG 291  
DB 121 AAGCTGCTGTGCTGCGCGCGGAGGAGGCTACCGCGGCGCGCTGCTG 180  
QY 292 TCCATGAGCGGCTGAGCTCAACGAAAGTGTCACTTCTGAGGAGACTTACGCGC 351  
DB 181 TCCATGAGCGGCTGAGCTCAACGAAAGTGTCTTCTGAGGAGACTTACGCGC 240  
QY 352 GCCGAGCTCACCGCTAAGCTGTGCGCGACATGGGCTGCAAGATGCCCGGACGCTG 411  
DB 241 GCCGAGCTCACCGCTAAGCTGTGCGCGACATGGGCTGCAAGATGGCCCGGACGCTG 300  
QY 412 CAGGCGGCCACGACAGGG 431  
DB 301 CAGGCGGCCACGACAGGG 320

RESULT 8 830 bp mRNA linear EST 14-NOV-2000  
BF238010 LOCUS 601811713F1 NIH\_MGC\_48 Homo sapiens CDNA clone IMAGE:4054731 5',  
DEFINITION mRNA sequence.

ACCESSION BF238010  
VERSION BF238010.1 GI:11151929  
KEYWORDS EST.  
SOURCE human.  
ORGANISM Homo sapiens

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.

REFERENCE 1 (bases 1 to 830)  
AUTHORS NIH-MGC <http://mgc.nci.nih.gov/>.  
TITLE National Institutes of Health, Mammalian Gene Collection (MGC)  
JOURNAL Unpublished (1999)  
COMMENT Contact: Robert Strausberg, Ph.D.

Email: cgaabs-r@mail.nih.gov  
Tissue Procurement: Louis M. Staedt, M.D., Ph.D.  
CDNA Library Preparation: Ling Hong/Rubin Laboratory  
CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)  
DNA Sequencing by: Incyte Genomics, Inc.  
Clone distribution: MGC clone distribution information can be  
found through the I.M.A.G.E. Consortium/LLNL at:  
<http://image.llnl.gov>  
Plate: LLM895 row: j column: 04  
High quality sequence stop: 670.

## FEATURES

## source

1. 830  
Location/Qualifiers

/organism="Homo sapiens"  
/db\_xref="taxon:9606"  
/clone="IMAGE:4054731"  
/clone\_lib="NIH\_MGC\_48"  
/tissue\_type="primary B-cells from tonsils (cell line)"  
/lab\_host="DH10B (phage-resistant)"  
/note="Organ: B-cells; Vector: pORF7; Site\_1: XhoI;  
Site\_2: EcoRI; CDNA made by oligo-dT priming.  
directionally cloned into EcoRI/XhoI sites using the  
following 5' adaptor: GGCACGAC(G). Size selected >500bp  
for average insert size 1.8kb. Library constructed by Ling  
Hong in the laboratory of Gerald M. Rubin (University of  
California, Berkeley) using ZAP-CDNA synthesis kit  
(Stratagene) and Superscript II RT (Life Technologies).  
Note: this is a NIH-MGC Library."

BASE COUNT 175 a 242 c 274 g 139 t  
ORIGIN

Query Match 50.6%; Score 317; DB 12; Length 830;

Best Local Similarity 100.0%; Pred. No. 2.7e-49;

Matches 317; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 115 GCACGGGGGTAGAGCGGCGCGGGGATCTGAGCCATGGGCGCGCGGAC 174  
DB 1 GCACGGGGGTAGAGCGGCGCGGGGATCTGAGCCATGGGCGCGCGGAC 60  
QY 175 GCCATCTCTGATGCGCTGAGAACCTGACCGCGGAGACTCAAGATTCAAGCTGAAG 234  
DB 61 GCCATCTCTGATGCGCTGAGAACCTGACCGCGGAGACTCAAGATTCAAGCTGAAG 120  
QY 235 CTGCTGTGCGTGGCGTGGCGGAGGCGGATCCCGGGGCGCGCTGTGTC 294  
DB 121 CTGCTGTGCGTGGCGTGGCGGAGGCGGATCCCGGGGCGCGCTGTGTC 180  
QY 295 ATGAGCGCTTGGACCTCAACGACAGCTGTCACTTCTACCTGGAAGCTTACGGCGCC 354  
DB 181 ATGAGCGCTTGGACCTCAACGACAGCTGTCACTTCTACCTGGAAGCTTACGGCGCC 240  
QY 355 GAGCTCACCGCTAAGCTGTGCGGACATGGGCTGCAAGATGGCGGACGCTGACG 414  
DB 241 GAGCTCACCGCTAAGCTGTGCGGACATGGGCTGCAAGATGGCGGACGCTGACG 300  
QY 415 GCGGCCACGACAGGG 431  
DB 301 GCGGCCACGACAGGG 317

RESULT 9 792 bp mRNA linear EST 20-FEB-2002  
BM549665 LOCUS 6544142 NIH\_MGC\_118 Homo sapiens CDNA clone IMAGE:5745741  
DEFINITION 5' mRNA sequence.

ACCESSION BM549665  
VERSION BM549665.1 GI:18785261  
KEYWORDS EST.  
SOURCE human.  
ORGANISM Homo sapiens

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.  
REFERENCE 1 (bases 1 to 792)  
AUTHORS NIH-MGC <http://mgc.nci.nih.gov/>.

TITLE National Institutes of Health, Mammalian Gene Collection (MGC)  
JOURNAL Unpublished (1999)  
COMMENT Contact: Robert Strausberg, Ph.D.  
Email: cgapbs-remail.nih.gov  
Tissue Procurement: Life Technologies, Inc.  
CDNA Library Preparation: Life Technologies, Inc.  
CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)  
DNA Sequencing by: Agencourt Bioscience Corporation  
Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LLNL at: http://image.llnl.gov  
Plate: LLM12769 row: c column: 22  
High quality sequence stop: 570.  
Location/Qualifiers  
1. 792  
/organism="Homo sapiens"  
/db\_xref="taxon:9606"  
/clone\_image="5745741"  
/clone\_lib="NIH\_MGC\_118"  
/tissue\_type="leukocyte"  
/lab\_host="DH10B"  
/note="Vector: pCMV-SPORT6; Site.1: NotI; Site.2: EcoRV (destroyed); RNA source: Leukocytes from anonymous pool of non-activated adult donors. Library is oligo-dT primed and directionally cloned (EcoRV site is destroyed upon cloning). Average insert size 1.7 kb, insert size range 1.2-3.3 kb. Library is normalized and enriched for full-length clones and was constructed by C. Gruber (Invitrogen). Research Genetics tracking code 027. Note: this is a NIH\_MGC Library."

BASE COUNT 182 a 241 c 242 g 122 t 5 others  
ORIGIN

Query Match 50.5%; Score 316; DB 13; Length 792;  
Best Local Similarity 100.0%; Pred. No. 4.2e-49;  
Matches 316; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 116 CAGCGGGGTGACGGCGGACGGCGGGGATCCTGAGACCTAGGGCGCGCGGACG 175  
|||||  
Db 26 CACGGGGGTGACGGCGGACGGCGGGGATCCTGAGACCTAGGGCGCGCGGACG 85  
|||||  
QY 176 CCAATCCGATGCGGGGAGAACTGACCGCGGAGAGCTCAAGAACTTCAAGCTGAAGC 235  
|||||  
Db 86 CCAATCCGATGCGGGGAGAACTGACCGCGGAGAGCTCAAGAACTTCAAGCTGAAGC 145  
|||||  
QY 236 TCGTGTGCGTGCCTGCGCGAGGCGCTACGCGGCGCTCCGCGGCGCTGCTGCA 295  
|||||  
Db 146 TCGTGTGCGTGCCTGCGCGAGGCGCTACGCGGCGCTCCGCGGCGCTGCTGCA 205  
|||||  
QY 296 TGGACGCTTGGACCTCAACGACAAGCTGGTCACTTCACTGAGAGACTAGGGCGCG 355  
|||||  
Db 206 TGGACGCTTGGACCTCAACGACAAGCTGGTCACTTCACTGAGAGACTAGGGCGCG 265  
|||||  
QY 356 AGCTCAACCCCTAAGCTGCTGCGGACATGAGGCGCTGAGAGATGCGCGGAGCTGACG 415  
|||||  
Db 266 AGCTCAACCCCTAAGCTGCTGCGGACATGAGGCGCTGAGAGATGCGCGGAGCTGACG 325  
|||||  
QY 416 CGGCGACGACCAAGG 431  
|||||  
Db 326 CGGCGACGACCAAGG 341  
|||||

RESULT 10  
BM926438 804 bp mRNA linear EST 12-MAR-2002  
LOCUS BM926438  
DEFINITION AGENCOURT 6644759 NIH\_MGC\_122 Homo sapiens cDNA clone IMAGE:5766806  
5', mRNA sequence.  
ACCESSION BM926438  
VERSION BM926438.1 GI:19376817  
KEYWORDS EST.  
SOURCE human.  
ORGANISM Homo sapiens  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

REFERENCE 1 (bases 1 to 804)  
AUTHORS NIH-MGC http://mgc.nci.nih.gov/.  
TITLE National Institutes of Health, Mammalian Gene Collection (MGC)  
JOURNAL Unpublished (1999)  
COMMENT Contact: Robert Strausberg, Ph.D.  
Email: cgapbs-remail.nih.gov  
Tissue Procurement: Life Technologies, Inc.  
CDNA Library Preparation: Life Technologies, Inc.  
CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)  
DNA Sequencing by: Agencourt Bioscience Corporation  
Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LLNL at: http://image.llnl.gov  
Plate: LLM12824 row: a column: 15  
High quality sequence stop: 746.  
Location/Qualifiers  
1. 804  
/organism="Homo sapiens"  
/db\_xref="taxon:9606"  
/clone\_image="5766806"  
/clone\_lib="NIH\_MGC\_122"  
/lab\_host="DH10B"  
/note="Organ: pooled lung and spleen; Vector: pCMV-SPORT6; Site.1: NotI; Site.2: EcoRV (destroyed); RNA source: anonymous pool of 24 week female lung, 16 week female spleen, and 20-22 week male spleens. Library is oligo-dT primed and directionally cloned (EcoRV site is destroyed upon cloning). Average insert size 1.4 kb, insert size range 1-3 kb. Library is normalized and enriched for full-length clones and was constructed by C. Gruber (Invitrogen). Research Genetics tracking code 026. Note: this is a NIH\_MGC Library."

BASE COUNT 165 a 238 c 243 g 125 t 33 others  
ORIGIN

Query Match 50.3%; Score 315; DB 14; Length 804;  
Best Local Similarity 100.0%; Pred. No. 6.5e-49;  
Matches 315; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 117 AGCGGGGTGACGGCGGACGGCGGGGATCCTGAGACCTAGGGCGCGCGGAGCG 176  
|||||  
Db 16 AGCGGGGTGACGGCGGACGGCGGGGATCCTGAGACCTAGGGCGCGCGGAGCG 75  
|||||  
QY 177 CATCTGATGCTGCTGAGAACTGACCGCGGAGAGCTCAAGAACTTCAAGCTGAAGCT 236  
|||||  
Db 76 CATCTGATGCTGCTGAGAACTGACCGCGGAGAGCTCAAGAACTTCAAGCTGAAGCT 135  
|||||  
QY 237 GCTGTGCGTGCCTGCGCGAGGCGCTACGCGGCGCTCCGCGGCGCGCTGCTGCCAT 296  
|||||  
Db 136 GCTGTGCGTGCCTGCGCGAGGCGCTACGCGGCGCTCCGCGGCGCGCTGCTGCCAT 195  
|||||  
QY 297 GGAGCGCTTGGACCTCAACGACAAGCTGCTTCACTTCACTGAGAGACTAGGGCGCG 356  
|||||  
Db 196 GGAGCGCTTGGACCTCAACGACAAGCTGCTTCACTTCACTGAGAGACTAGGGCGCG 255  
|||||  
QY 357 GCTCAACGCTAAGCTGCTGCGGACATGAGGCGCTGAGAGATGCGCGGAGCTGACG 416  
|||||  
Db 256 GCTCAACGCTAAGCTGCTGCGGACATGAGGCGCTGAGAGATGCGCGGAGCTGACG 315  
|||||  
QY 417 GGCGACGACCAAGG 431  
|||||  
Db 316 GGCGACGACCAAGG 330  
|||||

RESULT 11  
BG764161 750 bp mRNA linear EST 15-MAY-2001  
LOCUS BG764161  
DEFINITION 602737190F1 NIH\_MGC\_49 Homo sapiens cDNA clone IMAGE:4862182 5',  
mRNA sequence.  
ACCESSION BG764161  
VERSION BG764161.1 GI:14074814  
KEYWORDS EST.

SOURCE	ORGANISM	REFERENCE	AUTHORS	TITLE	JOURNAL	COMMENT
human.	Homo sapiens					
	Eukaryota: Metazoa: Chordata: Craniata: Vertebrata: Euteleostomi:					
	Mammalia: Eutheria: Primates: Catarrhini: Homnidae: Homo.					
	1 (bases 1 to 750)					
	NIH-MGC <a href="http://mgc.nci.nih.gov/">http://mgc.nci.nih.gov/</a> .					
	National Institutes of Health, Mammalian Gene Collection (MGC)					
	Unpublished (1999)					
	Contact: Robert Strausberg, Ph.D.					
	Email: <a href="mailto:cgapbs-r@mail.nih.gov">cgapbs-r@mail.nih.gov</a>					
	Tissue Procurement: ATCC/DCFD/DMP					
	cDNA Library Preparation: Ling Hong/Rubin Laboratory					
	cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)					
	DNA Sequencing by: Incyte Genomics, Inc.					
	Clone distribution: MGC clone distribution information can be					
	found through the I.M.A.G.E. Consortium/LLNL at:					
	<a href="http://image.llnl.gov">http://image.llnl.gov</a>					
	Plate: LHCMI722 row: d column: 23					
	High quality sequence stop: 721.					
FEATURES	location/Qualifiers					
source	1..750					
	/organism="Homo sapiens"					
	/db_xref="taxon:9606"					
	/clone="IMAGE:4862182"					
	/clone_lib="NIH-MGC_49"					
	/tissue_type="melanotic melanoma, high MDR (cell line)"					
	/lab_host="DH10B (phage-resistant)"					
	/note="Organ: skin; Vector: pORF8; Site.1: XhoI; Site.2:					
	EcoRI; cDNA made by oligo-dT priming. Directionally cloned					
	into EcoRI/XhoI sites using the following 5' adaptor:					
	GGGCAAG(G). Size-selected >500bp for average insert size					
	1.8kb. Library constructed by Ling Hong in the Laboratory					
	of Gerald M. Rubin (University of California, Berkeley)					
	using Zap-cDNA synthesis kit (Stratagene) and Superscript					
	II RT (Life Technologies). Note: this is a NIH-MGC					
	Library.  "					
BASE COUNT	165 a 232 c 235 g 118 t					
ORIGIN						
Query Match	50.2%; Score 314; DB 12; Length 750;					
Best Local Similarity	100.0%; Pred. No. 1e-48;					
Matches 314; Conservative	0; Mismatches 0; Indels 0; Gaps 0;					
Y 118	GGGGGGTAGGGGGGAGGGGGGAGGATCTTGAGCCATGGGGGCGCGCGAGAGCC 177					
Db 2	GGGGGGTAGGGGGGAGGGGGGAGGATCTTGAGCCATGGGGGCGCGCGAGAGCC 61					
Y 178	ATCTGTGATGCGCTGTGAGAACTGACCGCCGAGGAGGCTCAAGAAGTTCAAGCTGAAGCTG 237					
Db 62	ATCTGTGATGCGCTGTGAGAACTGACCGCCGAGGAGGCTCAAGAAGTTCAAGCTGAAGCTG 121					
Y 238	CTGTGTGATGCGCTGTGAGAACTGACCGCCGAGGAGGCTCAAGAAGTTCAAGCTGAAGCTG 297					
Db 122	CTGTGTGATGCGCTGTGAGAACTGACCGCCGAGGAGGCTCAAGAAGTTCAAGCTGAAGCTG 181					
Y 298	GAGCGCTTGAGACTTCACCGACAAGCTGTGACGCTTCTACCTGAGACCTACAGCGCCGAG 357					
Db 182	GAGCGCTTGAGACTTCACCGACAAGCTGTGACGCTTCTACCTGAGACCTACAGCGCCGAG 241					
Y 358	CTACAGCGTAACGTCGTGCGCGACATGGAGCTTCGAGAGAGATGGCGGAGCTTCAGAGCG 417					
Db 242	CTACAGCGTAACGTCGTGCGCGACATGGAGCTTCGAGAGAGATGGCGGAGCTTCAGAGCG 301					
Y 418	GCCACGACACGAGG 431					
Db 302	GCCACGACACGAGG 315					
RESULT 12						
BS6767422						
LOCUS	602767422					
DEFINITION	BCG74121991 NIH-MGC_49 Homo sapiens cDNA clone IMAGE:4871020 5',					
	mRNA sequence.					

Accession	Version	KeyWords	Source	Organism
B6767422	1	GI:14078075	human	Homo sapiens
B6767422.1	EST.			Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominiidae; Homo.
REFERENCE	AUTHORS	TITLE	JOURNAL	COMMENT
NH-MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LLNL at:	(bases 1 to 856)	NIH-MGC http://mgc.nci.nih.gov/	Unpublished (1999)	Contact: Robert Strausberg, Ph.D. Email: c9apbs-remail.nih.gov Tissue Procurement: ATCC/DCTP/DRP CDNA Library Preparation: Ling Hong/Rubin Laboratory DNA sequencing by: Incyte Genomics, Inc. Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LLNL at: http://image.llnl.gov Plate: LLCM1745 row e column: 05 High quality sequence stop: 696. Location/Qualifiers
FEATURES	SOURCE			
1..856	/organism="Homo sapiens"			
/db_xref="taxon:9606"				
/clone="IMAGE:4871020"				
/clone_id="NH_MGC_49"				
/tissue_type="melanotic melanoma, high MDR (cell line)"				
/lab_host="DH10B (phage-resistant)"				
/note="Organ: skin; Vector: pORF7; Site.1: XhoI; Site.2: EcoRI; CDNA made by oligo-dT priming. Directionally cloned into EcoRI/XhoI sites using the following 5' adaptor: GGACGACAG(G). Size-selected >500bp for average insert size 1.kbd. Library constructed by Ling Hong in the laboratory of Gerald M. Rubin (University of California, Berkeley) using ZAP-CDNA synthesis kit (Stratagene) and Superscript II RT (Life Technologies). Note: this is a NH_MGC Library."				
BASE COUNT	194 a	234 c	279 g	149 t
ORIGIN				
Query Match	50.2%; Score 314;	DB 12;	Length 856;	
Best Local Similarity	100.0%; Pred. No. 9,8e-49;			
Matches 314;	Conservative 0;	Mismatches 0;	Indels 0;	Gaps 0;
Y 118 GGGGGGTAGGGCGGAGCGCGGGGATCTTGAGCCATGTGGGGGGCGCGGGAGAGGCC 177				
D 2 GGGGGGTAGGGCGGAGCGCGGGGATCTTGAGCCATGTGGGGGGCGCGGGAGAGGCC 61				
Y 178 ATTCTGATGCGCTGAGGAACCTGAAACCGCGGAGAGAGCTCAAGAATTCAAAGTTAAGCTG 237				
D 62 ATCTGTGATGCGCTGAGGAACCTGAAACCGCGGAGAGAGCTCAAGAATTCAAAGTTAAGCTG 121				
Y 238 CTTCTGGTGGCCGCGCGGAGGCGCTTAAGGGGGGCAATCCCGGGGGGCGCGCTGTTCATG 297				
D 122 CTGTGGGTGGCCGCGCGGAGGCGCTTAAGGGGGCATCCCCGGGGGGCGCGTGTGTTCATG 181				
Y 298 GAGCGCTTGAGACTCACCGACAAGCTGTGATCAGCTTCTACTGTGAGAGACTTAAGGGCGCGAG 357				
D 182 GAGCGCTTGAGACTCACCGACAAGCTGTGATCAGCTTCTACTGTGAGAGACTTAAGGGCGCGAG 241				
Y 358 CTCACCGCTAACGTGTGGCGGAGCAATGGGCGCTGCAAGGAGATGTGCGGAGAGTGCGAGCG 417				
D 242 CTCACCGCTAACGTGTGGCGGAGCAATGGGCGCTGCAAGGAGATGTGCGGAGAGTGCGAGCG 301				
Y 418 GCCACGCAACGAGG 431				
D 302 GCCACGCAACGAGG 315				

LOCUS BG337806 748 bp mRNA linear EST 27-FEB-2001  
 DEFINITION 602435709F1 NIH\_MGC\_46 Homo sapiens cDNA clone IMAGE:4553378 5',  
 mRNA sequence.  
 ACCESSION BG337806  
 VERSION BG337806.1 GI:13144342  
 KEYWORDS EST.  
 SOURCE human.  
 ORGANISM Homo sapiens  
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.  
 REFERENCE 1 (bases 1 to 748)  
 NIH-MGC <http://mgs.nci.nih.gov/>.  
 National Institutes of Health, Mammalian Gene Collection (MGC)  
 Unpublished (1999)  
 JOURNAL  
 AUTHORS  
 TITLE  
 COMMENT  
 Contact: Robert Strausberg, Ph.D.  
 Email: [cgabs-remail.nih.gov](mailto:cgabs-remail.nih.gov)  
 Tissue Procurement: ATCC  
 CDNA Library Preparation: Ling Hong/Rubin Laboratory  
 CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LNL)  
 DNA Sequencing by: Incyte Genomics, Inc.  
 Clone distribution: MGC clone distribution information can be  
 found through the I.M.A.G.E. Consortium/LNL at:  
<http://image.llnl.gov>  
 Plate: LLM1251 row: b column: 03  
 High quality sequence stop: 710.  
 Location/Qualifiers  
 1. 748  
 /organism="Homo sapiens"  
 /db\_xref="taxon:9606"  
 /clone="IMAGE:4553378"  
 /clone\_lib="NIH\_MGC\_46"  
 /tissue="type="telomerosarcoma cell line"  
 /lab\_host="DH10B (phage-resistant)"  
 /note="Organ: uterus; Vector: pORF7; Site\_1: XhoI; Site\_2:  
 EcoRI; CDNA made by oligo-dT priming. Directionally cloned  
 into EcoRI/XhoI sites using the following 5' adaptor:  
 GGACGAG(G). Size selected >500bp for average insert size  
 1.8kb. Library constructed by Ling Hong in the laboratory  
 of Gerald M. Rubin (University of California, Berkeley)  
 using ZAP-cDNA synthesis kit (Stratagene) and Superscript  
 II RT (Life Technologies). Note: this is a NIH\_MGC  
 Library."  
 BASE COUNT 153 a 239 c 238 g 118 t  
 ORIGIN  
 Query Match 49.9%; Score 312.4; DB 12; Length 748;  
 Best Local Similarity 99.7%; Pred. No. 2e-48;  
 Matches 313; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

RESULT 14  
 LOCUS BM918855  
 DEFINITION AGENCOURT 6635188 NIH\_MGC\_120 Homo sapiens cDNA clone IMAGE:5747943  
 5', mRNA sequence.  
 ACCESSION BM918855  
 VERSION BM918855.1 GI:19369234  
 KEYWORDS EST.  
 SOURCE human.  
 ORGANISM Homo sapiens  
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.  
 REFERENCE 1 (bases 1 to 1294)  
 NIH-MGC <http://mgs.nci.nih.gov/>.  
 National Institutes of Health, Mammalian Gene Collection (MGC)  
 Unpublished (1999)  
 JOURNAL  
 AUTHORS  
 TITLE  
 COMMENT  
 Contact: Robert Strausberg, Ph.D.  
 Email: [cgabs-remail.nih.gov](mailto:cgabs-remail.nih.gov)  
 Tissue Procurement: Life Technologies, Inc.  
 CDNA Library Preparation: Life Technologies, Inc.  
 CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LNL)  
 DNA Sequencing by: Agencourt Bioscience Corporation  
 Clone distribution: MGC clone distribution information can be  
 found through the I.M.A.G.E. Consortium/LNL at:  
<http://image.llnl.gov>  
 Plate: LLM12774 row: o column: 16  
 High quality sequence stop: 304.  
 Location/Qualifiers  
 1. 1294  
 /organism="Homo sapiens"  
 /db\_xref="taxon:9606"  
 /clone="IMAGE:5747943"  
 /clone\_lib="NIH\_MGC\_120"  
 /lab\_host="DH10B"  
 /note="Organ: pooled pancreas and spleen; Vector:  
 pCW-SPORT6; Site\_1: NotI; Site\_2: EcoRV (destroyed); RNA  
 source anonymous pool of spleen and pancreas from 28 yo  
 male. Library is oligo-dT primed and directionally cloned  
 (EcoRV site is destroyed upon cloning). Average insert  
 size 1.5 kb, insert size range 1-2.5 kb. Library is  
 normalized and enriched for full-length clones and was  
 constructed by C. Gruber (Invitrogen). Research Genetics  
 tracking code 025. Note: this is a NIH\_MGC Library."  
 BASE COUNT 280 a 471 c 279 g 262 t  
 ORIGIN  
 Query Match 48.1%; Score 301.2; DB 14; Length 1294;  
 Best Local Similarity 97.5%; Pred. No. 2.2e-46;  
 Matches 306; Conservative 0; Mismatches 8; Indels 0; Gaps 0;



Db 312 TCAAGCATCATG 325

RESULT 15  
BM726665

LOCUS BM726665 446 bp mRNA linear EST 01-MAR-2002  
DEFINITION UI-E-EJ0-ali-n-17-0-UI.r1 UI-E-EJ0 Homo sapiens cDNA clone  
UI-E-EJ0-ali-n-17-0-UI 5', mRNA sequence.

ACCESSION BM726665  
KEYWORDS BM726665.1 GI:19047998

SOURCE

ORGANISM

human  
Homo sapiens  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Primates; Catarrhini; Homidae; Homo.

REFERENCE 1 (bases 1 to 446)  
AUTHORS Bernaldo,M.F., Lennon,G. and Soares,M.B.  
TITLE Normalization and subtraction: two approaches to facilitate gene  
discovery

JOURNAL Genome Res. 6 (9), 791-806 (1996)

COMMENT

Contact: Soares, MB  
Program for Rat Gene Discovery and Mapping  
University of Iowa  
451 Eckstein Medical Research Building Iowa City, IA 52242, USA  
Tel: 319 335 8250  
Fax: 319 335 9565  
Email: msoares@blue.weeg.uiowa.edu  
Tissue Procurement: Dr. Gregg Hageman  
cDNA Library preparation: Dr. M. Bento Soares, University of Iowa  
cDNA Library Arrayed by: Dr. M. Bento Soares, University of Iowa  
DNA Sequencing by: Dr. M. Bento Soares, University of Iowa  
Clone Distribution: Researchers may obtain clones from Research  
Genetics (www.resgen.com).  
Seq primer: M13 Reverse.

FEATURES  
source Location/Qualifiers

```
1..446
    /organism="Homo sapiens"
    /db_xref="taxon:9606"
    /clone="UI-E-EJ0-ali-n-17-0-UI"
    /clone_1lb="UI-E-EJ0"
    /tissue_type="fetal eyes, lens, eye anterior segment,
    optic nerve, retina, Retina Foveal and Macular, RPE and
    Choroid"
    /dev_stage="fetal and adult"
    /lab_host="DH10B (Life Technologies) (T1 phage resistant)"
    /note="Organ: eye; Vector: pT73-Pac (Pharmacia) with a
    modified polylinker; Site_1: EcoR I; Site_2: Not I;
    UI-E-EJ0 is a subtracted cDNA library constructed
    according to Bernaldo, Lennon and Soares, Genome Research,
    6:791-806, 1996. First strand cDNA synthesis was primed
    with an oligo-dT primer containing a Not I site. Double
    stranded cDNA was ligated to an EcoR I adaptor, digested
    with Not I, and cloned directionally into pT73-Pac
    vector. The oligonucleotide used to prime the synthesis of
    first-strand cDNA contains a library tag sequence that is
    located between the Not I site and the (dT)18 tail. The
    sequence tags for this library are: fetal eyes, AGAATCAAGA
    ; lens, CGATTAGCGA; eye anterior segment, AATGCCCAT;
    optic nerve, CCATTAGTG; retina, CCGCG; Retina Foveal and
    Macular, GTCC; RPE and Choroid, ACCCA. This library was
    created for the program, Gene Discovery in the Visual
    System, supported by National Eye Institute (NEI)."
```

BASE COUNT 76 a 145 c 156 g 69 t

ORIGIN

Query Match 47.3%; Score 296; DB 14; Length 446;  
Best Local Similarity 100.0%; Pred. No. 2.3e-45;  
Matches 296; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Oy 136 GGGGCGGGGATCCTGGAGCATGGGGCGGCGGAGCGGATCCTGGAGCGCTGGAG 195  
|||||  
Db 1 GGGGCGGGGATCCTGGAGCATGGGGCGGCGGAGCGGATCCTGGAGCGCTGGAG 60

```
Oy 196 AACCTGACCGCCGAGAGCTCAGAAAGTTCAAGCTGAAGCTGCTGTGCGGCGCTGGCC 255
|||||
Db 61 AACCTGACCGCCGAGAGCTCAGAAAGTTCAAGCTGAAGCTGCTGTGCGGCGCTGGCC 120
|||||
Oy 256 GAGGGCTACGGGCGCATCCCGCGGGCGGCTGCTGTCCATGACGCTTGGACCTCACC 315
|||||
Db 121 GAGGGCTACGGGCGCATCCCGCGGGCGGCTGCTGTCCATGACGCTTGGACCTCACC 180
|||||
Oy 316 GACAGCGTGTACGCTTCTACCTGAGAGCTTACGAGCTGAGCGGCGGAGCTTACGAGCTG 375
|||||
Db 181 GACAGCGTGTACGCTTCTACCTGAGAGCTTACGAGCTGAGCGGCGGAGCTTACGAGCTG 240
|||||
Oy 376 CGGACATGGGCTGTCAGAGAGATGGCGGGCAGCTGCAGCGGCGGACGACGAGG 431
|||||
Db 241 CGGACATGGGCTGTCAGAGAGATGGCGGGCAGCTGCAGCGGCGGACGACGAGG 296
|||||
```

Search completed: December 22, 2002, 06:45:56  
Job time : 1999 secs

